ADM
RARY

JUN 22 1943

Journal of the

AMERICAN STATISTICAL ASSOCIATION

CONTENTS *for* JUNE 1943

Factors in the Variability of Response in Enumerative Studies	*Gladys L. Palmer*	143
Time Series Significance Tests Based on Signs of Differences	*Geoffrey H. Moore and W. Allen Wallis*	153
An Approach to the Measurement of Farm Population Pressure in Wisconsin	*Thomas C. McCormick*	165
The Production Function for Canadian Manufactures	*Patricia Daly and Paul H. Douglas*	178
The Composition of Income and Ownership of Capital by Income Classes in the United States in 1936	*Rufus S. Tucker*	187
Comparison of Demographic Rates Assumed by the National Resources Committee with Actual Experience	*Robert J. Myers*	201
Choice of the Dependent Variable in Regression Analysis	*Frederick V. Waugh*	210
Comments	*Mordecai Ezekiel*	214
Note on a Common Statistical Inequality	*Irving H. Siegel*	217
Contents of Tippett's "Random Sampling Numbers"	*Robert Gage*	223
Quality Control Applied to Business Administration	*Wm. B. Rice*	228
Family Versus Individual in Measurement of Audiences	*Alfred Politz*	233
The Cost of Living in Cities in Relation to Their Size and Latitude	*E. L. Thorndike and Ella Woodyard*	238
Experience with Tests of Significance: A Reply to Professor R. A. Fisher	*Joseph Berkson*	242
Edward Lewis Dodd, 1875-1943	*C. D. Simmons*	247
The Study of Cycles		248
Book Reviews		250

VOLUME 38 NUMBER 222

American Statistical Association

Organized November 27, 1839
Incorporated 1841

The American Statistical Association is a scientific and educational organization. Its membership is not confined to professional statisticians but includes economists, business executives, research directors, government officials, university professors, and other persons who are seriously interested in the application of statistical methods to practical problems, in the development of more useful methods, and in the improvement of basic statistical data. Engineers, mathematicians, biologists, actuaries, sociologists, psychologists, and representatives of many other professions are included in the membership of the Association. Information about the Association and membership application forms may be secured from the Secretary, Lester S. Kellogg, American University, Washington, D. C.

War conditions have necessitated the indefinite postponement of the 104th Annual Meeting of the American Statistical Association which was to have been held in Cleveland, Ohio, December 29 to 31, 1942.

Chapters of the Association have been organized in important cities in the United States. Information concerning the meetings of these chapters may be secured from the District Representatives listed on another page.

The Editors welcome the submission of articles and notes for possible publication in the JOURNAL. Manuscripts should be sent to the Editor, Journal of the American Statistical Association, American University, Washington, D. C. Books for review should be sent to the Review Editor, Glenn E. McLaughlin, American University, Washington, D. C. Authors who wish suggestions about the preparation of manuscripts and charts or information about editorial policies should address their inquiries to the Editor.

Subscription rate, $6.00 per annum. Single numbers of volumes 23 to 38 $1.50 per copy postpaid, except Proceedings Supplements to volumes 23 to 30, $2.00 per copy postpaid. Prices for sets and discounts for orders in quantity available on request.

Published Quarterly by the AMERICAN STATISTICAL ASSOCIATION

Publication Office: 450 Ahnaip Street, Menasha, Wisconsin. Editorial Office: American University, Washington, D.C. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in paragraph 4, section 538, P. L. & R., authorized March 25, 1936. Entered as second class matter at the post office at Menasha, Wisconsin.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

JUNE, 1943 · $1.50 PER COPY · $6.00 PER ANNUM · VOL. 38 · NO. 222

CONTENTS

ARTICLES

Factors in the Variability of Response in Enumerative Studies . . . . GLADYS L. PALMER 143
Time Series Significance Tests Based on Signs of Differences . . . . GEOFFREY H. MOORE AND W. ALLEN WALLIS 153
An Approach to the Measurement of Farm Population Pressure in Wisconsin . . . . THOMAS C. MCCORMICK 165
The Production Function for Canadian Manufactures . . . . PATRICIA DALY AND PAUL H. DOUGLAS 178
The Composition of Income and Ownership of Capital by Income Classes in the United States in 1936 . . . . RUFUS S. TUCKER 187
Comparison of Demographic Rates Assumed by the National Resources Committee with Actual Experience . . . . ROBERT J. MYERS 201
Choice of the Dependent Variable in Regression Analysis . . . . FREDERICK V. WAUGH 210
Comments . . . . MORDECAI EZEKIEL 214
Note on a Common Statistical Inequality . . . IRVING H. SIEGEL 217
Contents of Tippett's "Random Sampling Numbers" . ROBERT GAGE 223
Quality Control Applied to Business Administration . . WM. B. RICE 228
Family Versus Individual in Measurement of Audiences ALFRED POLITZ 233
The Cost of Living in Cities in Relation to Their Size and Latitude . . . . E. L. THORNDIKE AND ELLA WOODYARD 238
Experience with Tests of Significance: A Reply to Professor R. A. Fisher . . . . JOSEPH BERKSON 242
Edward Lewis Dodd, 1875–1943 . . . . C. D. SIMMONS 247
The Study of Cycles 248
A Correction 249

BOOK REVIEWS

BEEBE, GILBERT WHEELER. *Contraception and Fertility in the Southern Appalachians.* Margaret Jarman Hagood 252
BENNETT, KENNETH R., see PEARSON, FRANK A.
BROOKINGS INSTITUTION. *Refugee Settlement in the Dominican Republic.* Erich W. Zimmermann 254
Commodity Statistics, The 1942 Commodity Year Book. Oscar L. Endler . . 260
Commodity Year Book. Oscar L. Endler 260
CULLITON, JAMES W. *Make or Buy.* Charles E. Young 261
DAVIS, HIRAM S. *Wool and the War.* R. R. Renne 262
GUILFORD, J. P. *Fundamental Statistics in Psychology and Education.* Martin A. Brumbaugh 266
HABERLER, GOTTFRIED. *Consumer Instalment Credit and Economic Fluctuations.* James F. Cusick 256

Continued on next page

HALL, MARGUERITE F. *Public Health Statistics.* Clyde V. Kiser 253
HOLMES, MAURICE C. *An Outline of Probability and Its Uses.* E. C. Molina 267
HORTON, DONALD C., LARSEN, HARALD C., and WALL, NORMAN J. *Farm-Mortgage Credit Facilities in the United States.* John B. Schneider . . 257
KISER, CLYDE V. *Group Differences in Urban Fertility.* Frank Lorimer . . 250
LARSEN, HARALD C., see HORTON, DONALD C.
LEAGUE OF NATIONS. *Wartime Rationing and Consumption.* M. K. Bennett 264
PAYNE, WILSON F. *Business Behavior 1919–1922.* V. P. Timoshenko . . 258
PEARSON, FRANK A., and BENNETT, KENNETH R. *Statistical Methods Applied to Agricultural Economics.* E. L. Butz 265
PETERS, J., compiled by. *Seven-Place Values of Trigonometric Functions.* Kenneth N. Ogle 268
WALL, NORMAN J., see HORTON, DONALD C.

PUBLICATIONS RECEIVED 269

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 222 JUNE, 1943 *Volume 38*

FACTORS IN THE VARIABILITY OF RESPONSE IN ENUMERATIVE STUDIES

By Gladys L. Palmer
Industrial Research Department, University of Pennsylvania

There are several sources of error in enumerative studies. The error that arises in sampling is well recognized, but other sources, which may be more important, are often overlooked, even though they are present as much in complete coverages as in samples. For instance, there is a variability of response to be reckoned with even within the framework of well defined schedule concepts and well trained enumerators. Poorly worded schedule questions and inadequate training of enumerators not only intensify this variability, but are the cause of unreliability in the sense of biases. However, in the absence of a standard operation, which by definition would yield no error and would enable one to differentiate correct from incorrect results, and thus to recognize biases, it is perhaps better to speak of variability of response merely as a source of difference between the results of two or more coverages of the same universe, rather than to speak of a bias in one or both coverages. The purpose of this paper is to present some results on variability of response.[1]

Relatively little is known about the degree of variability of response for different types of information or its relationship to other conditions of enumeration because the opportunity to measure it is limited to studies in which two sets of interviews requesting the same information for the same persons for the same period of time have been made. Relatively high costs have discouraged this kind of enumerative study while unfortunately, at the same time, differences in the results from applications of parallel techniques to the same universe for measuring the same phenomenon, without adequate explanation of the causes of these differences, have led to public confusion. As a consequence, doubt has been cast on the results of enumerative studies, yet these are the

[1] The writer is indebted to Dr. W. Edwards Deming and Dr. A. Ross Eckler of the Bureau of the Census for helpful suggestions with regard to this analysis.

only, or the major, source of information for many types of important social and economic data.[2]

An opportunity to study variability of response is found in two sets of interviews of 8519 persons in 2686 Philadelphia households obtained by two groups of enumerators who had been subject to the same training program. The purpose of this analysis is to determine the relation of certain conditions of enumeration to variation in individual response and the effect of variation in individual response upon the summarized results.

Four important social and economic characteristics were selected for the analysis of variability of response: marital status, age, employment status, and education. These characteristics were related to the following conditions of enumeration:

1. The proportion of Negro or foreign-born residents, or both combined.
2. The number of days of experience of enumerators at the time of the first field visit.
3. The lapse of time between the two interviews.
4. The personnel ratings of the original and revisit enumerators.[3]

Findings of this Study.—The findings of this study indicate that in the census type of survey, where the housewife is the usual respondent, one may expect a negligible variation response for marital status (1.7 per cent), a higher variation in response for employment status (9.9 per cent), and age (12.8 per cent), and the highest for education (23.4 per cent). These differences in variability of response arise partly from inherent difficulties in identification of status or classification. This is especially true with respect to returns on employment status for persons whose labor market attachments are difficult to identify or classify. It is also true for information on education for persons whose schooling has been informal or is described in terms of the educational systems of foreign countries which differ from the American school system. Variation in response for age or marital status cannot be explained in these terms, however, and is apparently the result of indifferent or inaccurate reporting.

Variability of response is lower for returns from districts in which there is a relatively high proportion of native-white residents than for

[2] For a summary of recent studies discussing this problem, see Gladys L. Palmer, *The Reliability of Response in Labor Market Inquiries*, Technical Paper #22, *Report to Federal Statistical Agencies*, Bureau of the Budget, July 1942.

[3] The personnel rating was based on all work of the enumerators during the survey and not on the results of the revisit program. Ratings were made by supervisors for personal appearance and personality, accuracy, learning ability, and output. These were averaged to give grades A, B, C, D, E, and F and yield a rough estimate of the comparative ranking of the enumerators' work.

districts in which there is a relatively high proportion of Negroes or of Negroes and foreign-born white residents combined.

None of the other conditions of enumeration which were selected for analysis in this study shows consistent relationships to variability of response for all four characteristics, nor are there significant differences in the variability of response for returns for men as compared with returns for women. Although individual responses in the two field visits show a relatively high variability for age and education, the summarized results for these items in the two enumerations in terms of frequency distributions and averages, are stable. This is not so, however, in the case of employment status.

Providing the enumerators measure up to a certain minimum standard, as the enumerators did for this survey, there is little correlation beyond that point between the individual's subjective rating and his performance as measured by variability of response. Variability of response is to be expected even under the best enumerating conditions, and this may reach surprisingly high proportions, as is indicated by the figures quoted above.

Source of Information.—There were 2686 Philadelphia households in the revisit sample, with a count of 8848 persons 16 years of age and over in the first enumeration and 8843 in the second.[4] There were 8519 persons for whom data could be matched for both visits, although all of the information analyzed in this study was not reported on both visits for each person enumerated twice.[5] The comparative data presented show differences for those of the total group of 8519 persons for whom the items were reported twice.

Selection of Enumerators.—The 36 enumerators employed in the revisit program were, on the whole, more experienced in schedule interviewing than the 197 enumerators who made the first visit. Both groups of enumerators, however, were given the same training with respect to the use of the schedule. Both groups were composed of high school or college graduates most of whom had had business, teaching, or social work experience. The information called for in the second visit referred to conditions as of the date of the first visit, and the second enumerators had no information about the household except the ad-

[4] These counts included 324 persons 16 years of age and over who were included in the second enumeration but not in the first, and 329 included in the first and not in the second. This difference in the count of persons 16 years of age and over arises to some extent from actual changes in household composition, residence, or age between the two visits and from differences in the application of the "household" concept by the two groups of enumerators, but is primarily the result of inconsistencies in the reporting of respondents.

[5] For persons enumerated in both visits, employment status was reported for all persons, marital status for all but a few individuals, age for 99 per cent, and education for 95.8 per cent. It may be noted that there were more failures to report when a different respondent gave the information on the second visit.

dress.[6] An average (median) of 8.7 days elapsed between the two visits, and all revisits had been completed within 23 days.

In 63 per cent of the households, the same person answered the questions on both visits of the enumerators. In about three-fifths of these cases, the respondent was the housewife. Household heads or other relatives were the respondents in the remainder of the cases in which the same respondent gave the information on both visits.

TABLE I

PER CENT REPORTING DIFFERENT RESPONSE BY CITY DISTRICT COMBINATIONS

City district combinations	All respondents			Same respondent			Different respondent		
	Total	Different		Total	Different		Total	Different	
	Number	Number	Per cent	Number	Number	Per cent	Number	Number	Per cent
Marital status									
Total*	8515	141	1.7	5373	87	1.6	3142	54	1.7
Districts 1, 7, 8, 9, 10	4293	58	1.4	2800	36	1.3	1493	22	1.5
Districts 2, 3, 4, 5, 6†	4222	83	2.0	2573	51	2.0	1649	32	1.9
Age									
Total*	8472	1086	12.8	5349	552	10.3	3123	534	17.1
Districts 1, 7, 8, 9, 10	4262	525	12.3	2786	270	9.7	1476	255	17.3
Districts 2, 3, 4, 5, 6†	4210	561	13.3	2563	282	11.0	1647	279	16.9
Employment status									
Total*	8519	843	9.9	5375	455	8.5	3144	388	12.3
Districts 1, 7, 8, 9, 10	4293	367	8.5	2800	204	7.3	1493	163	10.9
Districts 2, 3, 4, 5, 6†	4226	476	11.3	2575	251	9.7	1651	225	13.6
Education									
Total*	8165	1911	23.4	5204	1133	21.8	2961	778	26.3
Districts 1, 7, 8, 9, 10	4092	907	22.2	2706	552	20.4	1386	355	25.6
Districts 2, 3, 4, 5, 6†	4073	1004	24.7	2498	581	23.3	1575	423	26.9

* Excludes persons not reporting items tabulated.

† Districts with 40 per cent or more Negro and/or foreign-born residents.

Method of Analysis.—The results of a matching of the two sets of schedules on a household basis, as distinct from the persons composing the households, have already been reported in another publication.[7] The present analysis is made on the basis of a posting of the two sets of

[6] The 2686 addresses for the 1937 revisit program had been selected at random from a larger sample of 45,927 households in 149 selected School Blocks scattered throughout the city of Philadelphia, which had been used as a sample in a series of surveys of employment and unemployment from 1931 to 1938. The purpose of the check was to test the application of schedule concepts by the original enumerators as well as to check variability of response. The selection of revisit addresses was such that the work of all the original enumerators could be reviewed.

[7] This matching showed a relatively high variability of response for such items as age, employment status, occupation and industry, duration of jobs or unemployment periods, and years of residence in the city, as compared with such items as race, nativity, relationship to household head, and marital status. Margaret W. Beil and Gladys L. Palmer, *Employment and Unemployment in Philadelphia in 1936 and 1937*, Part 2, W.P.A. National Research Project in cooperation with the Industrial Research Department, University of Pennsylvania, 1938, Appendix A.

responses for each person in the sample households. The data were tabulated by the sex of the person reported, by whether the respondent was the same or different, and by city district.[8] The proportion of individuals for whom different information was reported on the second as compared with the first visit varied only slightly as between districts in the city. Variability of response was measured by the degree to which the returns for the two visits fell into different class-intervals when grouped.[9]

TABLE II

PER CENT REPORTING DIFFERENT RESPONSE BY NUMBER OF DAYS IN FIELD AT FIRST VISIT

Days in field at first visit	All respondents			Same respondent			Different respondent		
	Total	Different		Total	Different		Total	Different	
	Number	Number	Per cent	Number	Number	Per cent	Number	Number	Per cent
Marital status									
Total*	8484	140	1.7	5346	86	1.6	3138	54	1.7
First week	4631	77	1.7	2899	44	1.5	1732	33	1.9
Second week and later	3853	63	1.6	2447	42	1.7	1406	21	1.5
Age									
Total*	8442	1084	12.8	5323	550	10.3	3119	534	17.1
First week	4609	579	12.5	2890	298	10.3	1719	281	16.3
Second week and later	3833	505	13.2	2433	252	10.4	1400	253	18.1
Employment status									
Total*	8488	835	9.8	5348	448	8.4	3140	387	12.3
First week	4653	471	10.1	2901	248	8.6	1752	223	12.7
Second week and later	3835	364	9.5	2447	200	8.2	1388	164	11.8
Education									
Total*	8137	1905	23.4	5180	1127	21.8	2957	778	26.3
First week	4430	1053	23.8	2799	633	22.8	1631	420	25.7
Second week and later	3707	852	23.0	2381	494	20.7	1326	358	37.1

* Excludes persons not reporting items tabulated.

Factors in the Variability of Individual Response.—For all items except marital status, variability of response was significantly higher in cases in which the respondent was a different person on the two visits. The accompanying tables show this distinction. There seem to be no significant differences in the variability of response in reports for men as compared with women; consequently the data for the sexes have been combined in the analysis of differences in individual response.

Variability of response is found to be higher for districts with 40 per cent or more Negro or foreign-born residents or both combined, than for districts predominantly composed of native-born white residents

[8] The districting of the city follows the pattern of the School District divisions in use in 1937 when the original survey was made. All sections of the city were included.

[9] The class-intervals used for frequency distributions will be found in Table IV.

(Table I). The difference is not significant for returns for age, but is significant for returns for marital status, education, and employment status. It may be noted that this relationship holds for the returns for all four items when these were given by the same respondent, and for all items except age for returns given by a different respondent.

TABLE III

PER CENT REPORTING DIFFERENT RESPONSE BY DAYS BETWEEN TWO VISITS

Days between two visits	All respondents			Same respondent			Different respondent		
	Total	Different		Total	Different		Total	Different	
	Number	Number	Per cent	Number	Number	Per cent	Number	Number	Per cent
Marital status									
Total*	7394	123	1.7	4656	76	1.6	2738	47	1.7
1– 7	1544	25	1.6	976	15	1.5	568	10	1.8
8–14	5376	84	1.6	3417	54	1.6	1959	30	1.5
15–23	474	14	3.0	263	7	2.7	211	7	3.3
Age									
Total*	7354	944	12.8	4635	480	10.4	2719	464	17.1
1– 7	1533	197	12.9	971	89	9.2	562	108	19.2
8–14	5350	672	12.6	3402	359	10.6	1948	313	16.1
15–23	471	75	15.9	262	32	12.2	209	43	20.6
Employment status									
Total*	7395	735	9.9	4656	393	8.4	2739	342	12.5
1– 7	1544	150	9.7	976	68	7.0	568	82	14.4
8–14	5377	534	9.9	3417	303	8.9	1960	231	11.8
15–23	474	51	10.8	263	22	8.4	211	29	13.7
Education									
Total*	7085	1658	23.4	4509	976	21.6	2576	682	26.5
1– 7	1472	338	23.0	937	193	20.6	535	145	27.1
8–14	5159	1216	23.6	3319	728	21.9	1840	488	26.5
15–23	454	104	22.9	253	55	21.7	201	49	24.4

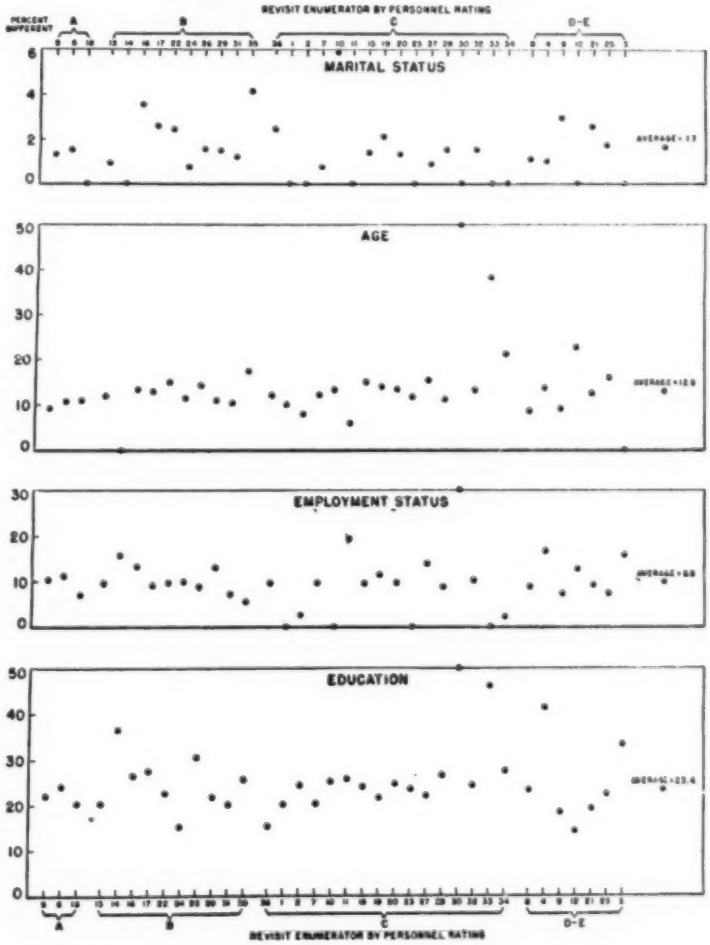
* Excludes persons not reporting items tabulated.

Compared with later returns, the results obtained during the first week of the first enumeration show only a slightly higher variability in the returns for marital status, employment status, and education, but not for age (Table II). It may be inferred that even well trained enumerators may become careless after the initial training wears off, and the effect of this carelessness almost counterbalances the steadying influence of experience. Moreover, the greater the lapse of time between the two visits, the greater the variability of response on marital status, age, and employment status (Table III), but not on education.

Variability Arising from Rating of Revisit Enumerators.—The revisit enumerators were rated subjectively A, B, C, D, E (footnote 3), and the accompanying chart was plotted to learn whether there is any relation between an individual's rating and variability of response, as

VARIATION IN RESPONSE FOR MARITAL STATUS, AGE, EMPLOYMENT STATUS, AND EDUCATION

By Personnel Rating of Revisit Enumerator



measured by the per cent of responses which differed in the two visits. There is a slight indication that A and B revisit enumerators showed less variability than those graded D and E. A few revisit enumerators showed variability outside the control limits (not drawn), but they are found in B ratings as well as D and E. This chart is consistent with the statement regarding the enumerators, that beyond a certain standard (surpassed by even the E revisit enumerators), there is little relation between performance as measured by variability of response and rating.

This does not mean that the general quality of schedule reporting is not better with experienced than with inexperienced enumerators, but that within any given framework of schedule concepts and training and qualifications of enumerators, variability of response is to be expected.

These findings support the evidence of another study in which variability of response in the work histories of skilled metal workers in Minneapolis was analyzed. In this study, many more factors were controlled than in a census type of survey, because a homogeneous group of workers was interviewed. Nevertheless, it was found that there was no consistent relationship between variability of response and the personality of interviewers or the lapse of time between the two interviews. Moreover, there was no consistent relationship between variability of response and a worker's age, education, or his nativity. The major factors which appeared to influence variability of response in this study of work history information were the simplicity or complexity of the data to be recalled and combinations of personal or psychological factors in response which could not be measured from available statistical data.[10]

Effects of Variation in Individual Response on the Summarized Results.—Although the degree of variability of response for individuals was relatively high for age and education, the effects of this are less significant when the data from the two enumerations are summarized (Table IV). The frequency distributions and averages secured for age and education show no significant differences, either for the total, or for men and women separately; nor are there significant differences in the proportions reporting various marital status categories in the two enumerations.

With regard to employment status, on the other hand, there are significant differences in the proportions reporting in certain categories in the two enumerations, when compared with each other and with the results of a combination of the two visits. The second visit yielded a larger labor force count than the first visit because it secured a considerably higher unemployment count. These differences are significant when related to the results of both visits combined. In the case of men, the count of the unemployed, and in the case of the women, the counts of the labor force, of non-workers, and of the unemployed differ significantly. This finding is corroborated by the results of other comparisons of employment status returns for the same individuals.[11]

The finding that, in general, summarized results are more stable than individual responses, is supported by other studies.[12] Many inconsist-

[10] Unpublished manuscript by Robert Dinkel in the files of the National Research Project of the Works Projects Administration.

[11] Palmer, *op. cit.*, footnote 2.

[12] Read Bain, "Stability in Questionnaire Response," *American Journal of Sociology*, November 1931, pp. 445–453; Ruth S. Cavan, "The Questionnaire in a Sociological Research Project," *American*

TABLE IV

MARITAL STATUS, AGE, EMPLOYMENT STATUS, AND EDUCATION BY SEX AS REPORTED IN THE FIRST AND SECOND ENUMERATIONS

	First enumeration						Second enumeration					
	Total		Men		Women		Total		Men		Women	
	Number	Per cent	Number	Per cent	Number	Per cent	Number	Per cent	Number	Per cent	Number	Per cent
Marital Status												
Total*	8846	100.0	4449	100.0	4397	100.0	8822	100.0	4431	100.0	4391	100.0
Single	3184	36.0	1775	39.9	1409	32.1	3213	36.4	1793	40.5	1420	32.3
Married	4631	52.3	2335	52.5	2296	52.2	4535	51.4	2279	51.4	2256	51.4
Other†	1031	11.7	339	7.6	692	15.7	1074	12.2	359	8.1	715	16.3
Age												
Total*	8834	100.0	4450	100.0	4384	100.0	8782	100.0	4419	100.0	4363	100.0
16–19	970	11.0	465	10.4	505	11.5	987	11.3	471	10.7	516	11.8
20–24	1241	14.0	648	14.6	593	13.5	1227	14.0	648	14.7	579	13.3
25–29	1022	11.6	501	11.3	521	11.9	1011	11.5	479	10.8	532	12.2
30–34	790	8.9	417	9.4	373	8.5	761	8.7	401	9.1	360	8.2
35–39	818	9.3	389	8.7	429	9.8	827	9.4	412	9.3	415	9.5
40–44	834	9.4	422	9.5	412	9.4	819	9.3	395	8.9	424	9.7
45–49	785	8.9	402	9.0	383	8.8	782	8.9	393	8.9	389	8.9
50–54	739	8.4	381	8.6	358	8.2	732	8.3	385	8.7	347	8.0
55–59	545	6.2	289	6.5	256	5.8	530	6.0	287	6.5	243	5.6
60–64	451	5.1	210	4.7	241	5.5	463	5.3	223	5.0	240	5.5
65 and over	639	7.2	326	7.3	313	7.1	643	7.3	325	7.4	318	7.3
Median age		37.4		37.5		37.3		37.5		37.6		37.3
Employment Status												
Total*	8519	100.0	4266	100.0	4253	100.0	8519	100.0	4266	100.0	4253	100.0
Labor force	5444	63.9	3858	90.4	1586	37.3	5569	65.4	3891	91.2	1678	39.5
Employed	3312	38.9	2317	54.3	995	23.4	3222	37.8	2249	52.7	973	22.9
Unemployed	1797	21.1	1268	29.7	529	12.4	2017	23.7	1374	32.2	643	15.1
Emergency	335	3.9	273	6.4	62	1.5	330	3.9	268	6.3	62	1.5
Not seeking work	3075	36.1	408	9.6	2667	62.7	2950	34.6	375	8.8	2575	60.5
Education												
Total*	8442	100.0	4218	100.0	4224	100.0	8207	100.0	4069	100.0	4138	100.0
0	483	5.7	215	5.1	268	6.4	470	5.7	194	4.8	276	6.7
1– 4	747	8.9	374	8.9	373	8.8	714	8.7	370	9.1	344	8.3
5– 8	4206	49.8	2129	50.5	2077	49.2	4180	51.0	2121	52.1	2059	49.8
9–12	2651	31.4	1284	30.4	1367	32.4	2501	30.5	1176	28.9	1325	32.0
13–16	318	3.8	185	4.4	133	3.1	290	3.5	164	4.0	126	3.0
17 and over	37	0.4	31	0.7	6	0.1	52	0.6	44	1.1	8	0.2
Median		7.8		7.9		7.8		7.8		7.8		7.8

* Excludes persons not reporting items tabulated.

† Widowed, separated, divorced.

encies in reporting are inevitably compensating in character when the data are combined into frequency distributions. The extent to which variability in individual responses affects cross-classifications, such as employment status by age and sex or age and sex by education, has not been explored here but the chances of variability in the summarized results of such cross-tabulations are obviously greater.

It is clear from this study that a complete census of any group in the population is subject to some error or variability in reporting, regardless of the perfection of schedule concepts and definitions or the training of enumerators. It is not necessary to emphasize that variability of response affects the analysis of all population data, and is also important in the development of sampling techniques for enumerative studies. The choice of economic or social controls for stratification of sampling and sub-sampling units must be determined, in part, by the relative reliability of the information obtained in the survey used for controls.

Journal of Sociology, March 1933, pp. 721–727; Alberta Curtis, "The Reliability of a Report on Listening Habits," *Journal of Applied Psychology*, Volume 23, 1939, pp. 127–130; Mapheus Smith, "A Note on Stability in Questionnaire Response," *American Journal of Sociology*, March 1933, pp. 713–727.

TIME SERIES SIGNIFICANCE TESTS BASED ON SIGNS OF DIFFERENCES*

By Geoffrey H. Moore
National Bureau of Economic Research
and W. Allen Wallis
Stanford University

Recently we proposed a significance test, appropriate to time series, based on signs of first differences in either the observations or some derivative sequence.[1] The test is made by tabulating the frequency distribution of phase durations—in other words the frequency distribution of runs in signs of first differences—and comparing it with the distribution expected if the series were random, $\frac{5(N-3)}{12}$, $\frac{11(N-4)}{60}$, and $\frac{4N-21}{60}$ being the expected frequencies for phases of durations one, two, and over two, respectively. Since this expected distribution is the same for all continuous populations, such a test has a high degree of generality. The ideal statistic for the comparison is not known, but for practical purposes it suffices to calculate an ordinary χ^2 for goodness of fit (denoting it by χ_p^2) and refer 6/7 of its value to the usual χ^2 table for two degrees of freedom when $\chi_p^2 < 6.3$, or refer its whole value to a special tabulation of χ^2 for $2\frac{1}{2}$ degrees of freedom when $\chi_p^2 > 6.3$.

A variation upon the χ_p^2 test is to tabulate separate frequency distributions of durations of "expansions" (runs up) and "contractions" (runs down). Each distribution may be compared with an expected distribution having frequencies half as large as those expected for the combined distribution, or a test of homogeneity may be applied to the two distributions. Theoretical difficulties attach to either alternative, however.[2]

An alternative (but not independent) test can be made by simply counting the number of phases, i.e., the number of runs in signs of differences.[3] This is equivalent to counting the number of turning points, i.e., reversals in direction of movement. Unless N, the number of ob-

* This paper is a byproduct of work at the National Bureau of Economic Research, 1939–40, under Research Associateships provided by the Carnegie Corporation of New York.

[1] W. Allen Wallis and Geoffrey H. Moore, *A Significance Test for Time Series and Other Ordered Observations*. National Bureau of Economic Research *Technical Paper 1* (September, 1941), 59 pp. For a brief summary of this monograph, see W. Allen Wallis and Geoffrey H. Moore, "A Significance Test for Time Series Analysis," this Journal, 36 (1941), pp. 401–9.

[2] See the *Technical Paper*, *op. cit.*, pp. 40–41. In addition to the difficulties there mentioned, it seems doubtful that the χ_p^2 distribution is appropriate to chi-squares derived from separate distributions of expansions and contractions.

[3] *Op. cit.*, Sec. VII.

servations, is small (in which case tabulations of the exact distributions are available), such counts are normally distributed with variance of $\frac{16N-29}{90}$, and mean of $\frac{2N-4}{3}$ for the number of turning points or $\frac{2N-7}{3}$ for the number of phases.

The present paper discusses three additional tests based on signs of differences:

(1) The total number of signs of each kind may be counted. The exact distributions are tabulated in Table I for $N \leqq 12$, and for longer series the number of signs of a given kind is normally distributed about $\frac{N-1}{2}$ with variance of $\frac{N+1}{12}$.

(2) The regularity of a periodic pattern, as judged by the tendency of the signs of one kind to cluster together in the pattern, can be tested, making allowance for any excess of signs of one kind. The exact probabilities can be calculated from a recursion formula, but have been tabulated only for the particular case of a nine point pattern.

(3) The existence of correlation between time series, abstracting from the effect of their trends, may be studied by making a 2×2 table of the signs of differences in the two series. The method may also be extended to more than two series. Although such a test would be of great value, the sampling distribution is known only for the special case of randomly arranged first differences.

In conclusion, a few general considerations relating to tests based on signs of differences are discussed.

I

A quick test of the randomness of a series can be made by counting the plus and minus signs in the first differences. A series of N observations has $N-1$ differences, and if the observations are in random order the expected number of plus or minus signs is obviously $\frac{N-1}{2}$. The variance of the number of signs of one kind is $\frac{N+1}{12}$, and the distribution rapidly approaches normality as N increases.

These results were reached by working with exact distributions for $N \leqq 12$, the exact distributions being calculated from the recursion relation

$$\phi_N(m) = (N - m)\phi_{N-1}(m - 1) + (m + 1)\phi_{N-1}(m),$$

where $\phi_N(m)$ represents the number of permutations of N different

numbers that produce exactly m positive (or m negative) differences.[4] For a series of values of N the variances were found to have constant first differences of $1/12$; since when $N=2$ the variance is $\frac{1}{4}$, the general expression[5] for the variance is $\frac{N+1}{12}$. The third moment about the mean is zero for all values of N, since the distributions are symmetrical. A series of values of the fourth moment about the mean shows, when $N \geqq 4$, constant second differences of $1/24$; since when $N=4$ the fourth moment about the mean is $23/48$ and the first difference $53/240$, the general expression for the fourth moment about the mean is $\frac{5N^2+8N+3}{240}$.

As N increases, the kurtosis, measured by the ratio of the fourth moment about the mean to the squared variance, approaches 3, the value for a normal distribution. For all values of N the skewness is, of course, zero, as it is for a normal distribution. Comparison of the exact distribution for $N=12$ with that based on a normal approximation shows fair agreement; so for $N>12$ the normal distribution apparently provides a satisfactory approximation. When using a normal approximation the Laplace correction for continuity, which consists of subtracting $\frac{1}{2}$ from the discrepancy between observed and expected frequencies, should be applied. (A formula for calculating the normal deviate is given in the note to Table I.) For $N \leqq 12$ the exact distributions are shown in Table 1.

A test of the number of signs of one kind may serve as a test of correlation between two variables. The paired values are arranged in order according to one variable, and the signs of differences are tabulated for the other variables. Such a test is perhaps most appropriate for the detection of trend—the correlation of a series with time—for here no question arises as to which of the variables is to be arranged in order. In other cases the choice may not be obvious, and the result may depend upon which variable is selected.

[4] This relation was derived empirically from equation 4, page 19, of the *Technical Paper*. As is pointed out there (pages 5 and 11), a distribution of signs of differences is exactly the same for random arrangements of N different numbers as for samples of N drawn from any continuous population.

After this paper was in proof we discovered that exactly the same recursion relation given above was given in a different context by Major P. A. MacMahon, "Second Memoir on the Compositions of Numbers," *Philosophical Transactions of the Royal Society of London*, 207A (1908), pp. 65–134. The recursion relation given above appears in Art. 30, p. 83. At the end of Art. 33, p. 81 of the same memoir, MacMahon gives an explicit formula which, in our notation, may be written

$$\phi_N(m)=\sum_{i=0}^{m}(-1)^i\binom{N+1}{i}(m+1-i)^N.$$

He also tabulates $\phi_N(m)$ through $N=6$ on page 81.

[5] This expression may also be deduced easily from the familiar theorem that the variance of the sum of k variates is the sum of the k variances plus twice the sum of the $[k(k-1)]/2$ covariances.

TABLE I

EXACT PROBABILITY THAT THE DISCREPANCY BETWEEN OBSERVED AND EXPECTED NUMBERS OF MINUS SIGNS IN THE FIRST DIFFERENCES OF A RANDOM SERIES OF N OBSERVATIONS WILL EQUAL OR EXCEED (*IN ABSOLUTE VALUE*) THE DISCREPANCY CORRESPONDING WITH ANY OBSERVED NUMBER OF MINUS SIGNS, TWO TO TWELVE OBSERVATIONS; AND NORMAL APPROXIMATION FOR TWELVE OBSERVATIONS

m*	N=2	N=3	N=4	N=5	N=6	N=7	N=8	N=9	N=10	N=11	N=12		m
											Exact	Approximate	
0	1.000,000	0.333,333	0.083,333	0.016,667	0.002,778	0.000,397	0.000,050	0.000,006	0.000,001	0.000,000	0.000,000	0.000,00	0
1	1.000,000	1.000,000	1.000,000	0.450,000	0.161,111	0.048,016	0.012,302	0.002,772	0.000,559	0.000,102	0.000,017	0.000,12	1
2		0.333,333	1.000,000	1.000,000	1.000,000	0.520,635	0.225,248	0.083,284	0.026,926	0.007,750	0.002,014	0.003.95	2
3			0.083,333	0.450,000	1.000,000	1.000,000	1.000,000	0.569,582	0.277,803	0.118,154	0.044,551	0.054,67	3
4				0.016,667	0.161,111	0.520,635	1.000,000	1.000,000	1.000,000	0.606,074	0.321,454	0.336,67	4
5					0.002,778	0.048,016	0.225,248	0.569,582	1.000,000	1.000,000	1.000,000	1.000,00	5
6						0.000,397	0.012,302	0.083,284	0.277,803	0.606,074	1.000,000	1.000,00	6
7							0.000,050	0.002,772	0.026,926	0.118,154	0.321,454	0.336,67	7
8								0.000,006	0.000,559	0.007,750	0.044,551	0.054,67	8
9									0.000,001	0.000,102	0.002,014	0.003,95	9
10										0.000,000	0.000,017	0.000,12	10
11											0.000,000	0.000,00	11

* m is the number of minus signs.

NOTE: For larger values of N, $\pm \dfrac{\sqrt{3}(|2m-N+1|-1)}{\sqrt{N+1}}$ may be regarded as normally distributed with zero mean and unit variance.

As an illustration, this test of the sign counts has been applied to sweet potato production, yield per acre, and acreage harvested in the United States,[6] 1868–1937. The production series shows 37 rises and 32 declines. Since $N=70$ (observe that N is the number of observations, not differences), the expected number of rises or declines is 34.5, which differs from the observed number by 2.5, or 2.0 when corrected for continuity. The variance is 71/12, or the standard deviation 2.432. Taking $2.0/2.432=0.822$ as a normal deviate, we find a probability of 0.41 for so large a departure from expectation. Hence there is no evidence of trend. For acreage, the probability is about 0.0000000001, clearly indicating the presence of trend.[7] Actually, the charts of the two series suggest that approximately the same degree of trend is present in both series. The production series, however, contains random movements whose year-to-year variations are much greater than the yearly trend effects; whereas in the acreage series, particularly in the earlier years, the random variations are small relative to the trend. Any test of significance judges systematic variations by the relation of their magnitude to that of random variations; in this case the consequence is detecting the trend in the more stable acreage series and missing a similar trend in the more erratic production series. (This point is discussed further in the final section.) The yield series shows 35.5 rises and 33.5 falls (counting the one year of no change as half a unit in each category). Since as great a difference would result 84 per cent of the time from random causes alone, there is no evidence of trend; and this accords with the appearance of the chart.

It should be emphasized that the null hypothesis is that the *observations* are in random order, not that the first differences are random. To illustrate the distinction, suppose that of the six differences in a series of seven observations, 5 are of one sign and 1 of the other. On the null hypothesis that the seven numbers are in random order, we find the exact probability of a division as unequal as five-to-one to be 0.048 (Table I), so should probably conclude that the discrepancy is significant. Had we, on the other hand, formulated the null hypothesis that the *differences* are random and equally likely to be positive or negative we should have used a binomial distribution and found a probability of 0.219, more than $4\frac{1}{2}$ times as great and definitely not significant. The variance on the null hypothesis that the differences are

[6] The original data (taken from *Agricultural Statistics, 1939*, p. 243) and a chart are shown in the *Technical Paper*, pp. 37–39.

[7] Similar results were obtained in the *Technical Paper*, pp. 40–41, by testing separate frequency distributions of expansion and contraction phases. The number of signs of one kind is of course equivalent to the total duration of phases of one kind, hence is related to their frequency distribution by duration. The test of sign count or total duration, however, is free from the theoretical difficulties that attach to a test of the frequency distributions.

random and equally likely to be positive or negative is $(N-1)/4$, and therefore (when $N>2$) exceeds the variance on the null hypothesis that the observations are random; as N increases, the ratio between the two variances approaches 3. Hence the first hypothesis yields larger probabilities for given departures from expectation.

In economic research it is not always easy, but it is of course essential, to determine which of the two null hypotheses is appropriate. In testing the randomness of changes in stocks of commodities, for example, it might be appropriate to use a binomial distribution in connection with signs of first differences of stocks; but in testing whether stocks themselves constitute a random series, the appropriate distribution is the one given here. It may be observed that an inequality in the sign counts which cannot be accounted for on the assumption of random differences certainly cannot be accounted for on the assumption of random observations; and an inequality that can be explained by the hypothesis of random observations can be still more easily explained by the hypothesis of random differences.

II

A second kind of problem for which the signs of differences provide an appropriate significance test arises in periodogram analysis and similar procedures. Suppose a series has been divided into equal segments, or segments derived according to some external criterion, e.g., the turning dates of cycles in general business. It may be desired to test whether the movements within such segments tend to be smooth and regular rather than zig-zags, since regularity can normally be taken as indirect evidence that the period used is reflected in the data. Perhaps a description of the particular situation that resulted in this application will clarify the problem.

One feature of the method of studying cyclical behavior devised by the National Bureau of Economic Research involves dividing time series into segments corresponding with the dates of so-called reference cycles, i.e., cycles in general economic activity, each segment being bounded by two consecutive reference troughs.[8] Within each reference cycle, the series is expressed in reference cycle relatives; that is, each observation is divided by the average value of the series during that reference cycle. Each reference cycle is then divided into nine stages, and the mean of the relatives within each stage is computed. The first and ninth stages each consist of three months centered on the reference troughs and the fifth stage consists of three months centered on the

[8] Cf. Wesley C. Mitchell and Arthur F. Burns, "The National Bureau's Measures of Cyclical Behavior," National Bureau of Economic Research *Bulletin 57* (July 1, 1935), 20 pp.; or better, their forthcoming volume, *Methods of Measuring Cyclical Behavior*.

reference peak; stages II, III, and IV correspond with successive thirds of the reference expansion, stages VI, VII, and VIII with successive thirds of the reference contraction. The nine means thus depict the pattern of behavior during a reference cycle.

Now if the series being analyzed contains cycles, any one reference pattern will be fairly regular, though its peaks and troughs need not conform to the reference dates. But an average reference cycle pattern made by averaging the reference stage standings over a number of reference cycles will tend to be simply a random sequence, unless the reference dates bear some relation to the behavior of the series—unless, in other words, the series conforms with general business activity.

The nine points of an average reference cycle pattern are rather few for a satisfactory test of the distribution of phase durations or of the total number of reversals. Furthermore, both these tests or one based on the total sign count would reflect the intra-cycle trend, which is retained in the reference cycle patterns. But from our recursion formula[9] it is possible to answer the question: Given the number of plus and minus signs in the average pattern, do they tend to cluster more than would be expected if they had occurred in a set of nine random standings?

To answer this question, consider the integers 1 to 8 arranged around a circle, with the signs of their first differences entered between them. The circular arrangement allows for the fact that stage IX in one cycle is based on the same data as stage I in the next, and eliminates the arbitrariness of taking up the cycle at any particular stage. Since 8 is the largest of the integers, the difference to one side of it must be positive and to the other side negative. From the recursion formula, the probability of each of the 64 arrangements of the other six signs that can occur among the 5,040 permutations of the digits 1 to 7 can be computed. Actually, we need consider only those of the 64 arrangements of signs that produce the observed count of pluses and minuses; if there are m minus signs ($1 \leqq m \leqq 7$), there are as many of these as there are combinations of six things taken $m-1$ at a time. The relevant arrangements of signs are then grouped according to the amount of clustering, as in Table II. The probability of as much clustering as is shown by each possible grouping is obtained by summing the probabilities of sequences of signs corresponding with each group.

This test is in a sense an amalgam of the test of phase durations and the test of turning points. Except for the circular arrangement, it is like the test of turning points in that the fewer the turning points the greater the clustering. But for a given number of turning points, the de-

[9] *Technical Paper*, Equation 4, p. 19.

TABLE II

EXACT PROBABILITY OF AS MUCH (OR MORE) REGULARITY OF PATTERN FROM EIGHT OBSERVATIONS ARRANGED AT RANDOM IN A CLOSED SEQUENCE AS IS SHOWN BY EACH POSSIBLE GROUPING OF SIGNS OF DIFFERENCES

4:4 Ratio between signs		5:3 Ratio between signs		6:2 Ratio between signs	
Grouping*	Probability	Grouping*	Probability	Grouping*	Probability
4 4	.0083	5 3	.0126	6 2	.0500
3–1 3–1	.0993	4–1 2–1	.1453	5–1 1–1	.2667
3–1 2–2	.2368	3–2 2–1	.3955	4–2 1–1	.7167
2–2 2–2	.2914	3–1–1 1–1–1	.6448	3–3 1–1	1.0000
2–1–1 2–1–1	.8874	2–2–1 1–1–1	1.0000		
1–1–1–1 1–1–1–1	1.0000				

* Entries under "grouping" indicate in the first line the lengths of runs in one sign, in the second line the lengths of runs in the other sign. Thus $\begin{matrix}3\text{–}1\\2\text{–}2\end{matrix}$ indicates that the pluses (minuses) are contained in one group of 3 and one group of 1 and the minuses (pluses) in two groups of 2 each.

gree of clustering is greater when the signs of one kind fall mostly in a single group (phase) than when they are more evenly distributed among the groups.

III

The test of cyclical conformity just described is based simply upon the regularity of the average pattern. Some types of systematic movement, e.g., seasonal, do not imply this type of regularity—indeed, it is possible that a series might follow business cycles systematically, but in an irregular fashion. This suggests the need for another kind of test, one to determine, for example, whether movements during corresponding stages of successive reference cycles are consistent. Such a test would be one of a general class of correlation tests based on signs of first differences.

Suppose it is desired to test whether there is a relation between the nine-point patterns of two reference cycles, the seasonal patterns given by twelve monthly figures for each of two years, or any two time series of equal length. A 2×2 table can be constructed by listing the signs of differences for each series and tabulating the number of instances in which (a) both series rise, (b) the first rises and the second falls, (c) the first falls and the second rises, and (d) both fall. A homogeneity test of

this double dichotomy, treating the marginal totals as fixed, would have the advantages of simplicity, generality, and elimination of at least the primary effects of trend.

The usual procedures for analyzing double dichotomies are based on the assumption that the entries are random and independent. If each series is a random sequence, the differences are negatively correlated serially, and this means that a chance coincidence in direction of movement during one interval tends to produce another coincidence during the next interval. Furthermore, suppose the first series has an excess of positive differences; these are more likely to be located at certain parts of the series than at others.[10] If the second series also has more plus than minus signs they tend to fall in the same positions, thus producing coincidences even though the series are unrelated; but if the second series has more minus than plus signs, the minuses tend to fall where the pluses do for the first series, thus producing a spurious appearance of a negative relation between the two series.

A few tabulations that we have made of the exact distribution on the assumption that the observations are random suggest that the usual

[10] This may seem surprising since each of the $N-1$ differences in a random series is equally likely to be positive or negative. When it is known, however, that one sign has occurred less frequently than the other, the occurrences of the less frequent sign are more likely to be at the center than at the extremes of the series. This may be verified by the recursion formula referred to in footnote 9, but the following example illustrates the point: Suppose that of the $N-1$ differences in a series of N, exactly one is negative. If this negative difference is the mth in the series of $N-1$ differences, it means that (1) the first m observations are arranged in ascending order, (2) the last $N-m$ observations are arranged in ascending order, and (3) the greatest observation among the first m exceeds the least among the last $N-m$. Any partition of the N observations into two groups, of m and $N-m$ respectively, provides for one (and only one) fulfillment of the first two conditions; so there are

$$\frac{N!}{m!(N-m)!}$$

ways to satisfy conditions 1 and 2. Of these, all but one—that in which each observation among the $N-m$ exceeds each among the m—also satisfy the third requirement. Hence there are among the permutations of N different numbers exactly

$$\frac{N!}{m!(N-m)!}-1$$

which have all differences positive except the mth. This quantity summed over the range $m=1$ to $m=N-1$, or 2^N-N-1, is the number of permutations of N different quantities that produce exactly one negative sign among the signs of differences. So if a random sequence of N numbers has produced only one negative first difference (the same reasoning applies for one positive difference), the probability that it is the mth difference is

$$\frac{N! - m!(N-m)!}{m!(N-m)!(2^N - N - 1)},$$

which has a maximum at $m=N/2$ when N is even or maxima at $m=(N-1)/2$ and $m=(N+1)/2$ when N is odd, and minima at $m=1$ and $m=N-1$. When there is more than one minus sign, but fewer than $(N-1)/2$, the situation is more complex. For example, two minus signs if adjacent are most likely to be in the center but if non-adjacent are most likely to be about one-third of the way from each end; they are more likely to be non-adjacent than to be adjacent and the net result appears to be a distribution with three modes, one at the center and the other two a little more than a third of the way from each end. As the number of minus signs rises toward $(N-1)/2$ the probability of a minus in any given position becomes more nearly uniform for all positions.

treatment of double dichotomies, though based on the assumption of random differences, may provide usable approximations unless N is very small or the number of plus and minus signs within either series is extremely unequal. When $N=8$ and each series has the same number of signs of a given kind, the comparison is as shown in Table III.

TABLE III

PROBABILITIES OF VARIOUS NUMBERS OF COINCIDENCES IN POSITION FOR SIGNS OF DIFFERENCES IN TWO RANDOM ARRANGEMENTS OF EIGHT OBSERVATIONS, EACH CONTAINING m POSITIVE (OR m NEGATIVE) DIFFERENCES; AND COMPARISON WITH APPROXIMATION BASED ON ASSUMPTION THAT DIFFERENCES ARE RANDOM AND INDEPENDENT

m*	Coincidences†	Probability	
		Exact	Approximate‡
1	1	0.202708	0.142857
	0	0.797292	0.857143
2	2	0.070409	0.047619
	1	0.470286	0.476190
	0	0.459305	0.476190
3	3	0.042049	0.028571
	2	0.342878	0.342857
	1	0.480747	0.514286
	0	0.134325	0.114286

* m is the frequency of the less frequent sign.

† Number of coincidences in position for the less frequent sign. The table also covers pairs of series in which one has m signs of one kind and the other has m signs of the opposite kind, if coincidences are counted for the two sets of m; in such cases inverse relationship is tested.

‡ Calculated by the method given in R. A. Fisher, *Statistical Methods for Research Workers*, Fifth and Later Editions (London and Edinburgh, 1934–40), Section 21.02.

It may be remarked that if the two dichotomies appear to be homogeneous when treated by the usual procedure, which implicitly assumes random differences, they surely would appear so if the proper null hypothesis of random observations were used; but nothing more can be said definitely until the sampling distribution appropriate to the assumption of random observations is known.

It is easy to extend this method of testing correlation to more than two series. For example, if it is desired to test the consistency of reference cycle behavior over several reference cycles, the signs of the differences for the eight stage-to-stage movements in the separate reference cycle patterns may be entered in a table of eight columns and as many rows as there are cycles. By counting the number of plus signs in each column, and comparing the counts with those expected if the total number of plus signs in the entire table were equally divided among the columns, we can see whether the movements during each interval tend

to be consistent from cycle to cycle, and at the same time allow for the effect of trend on the total number of plus signs. In testing the significance of this tendency, however, difficulties with respect to the null hypothesis again rise. If the differences can be assumed random sequentially, the variance of the column counts is $\frac{P}{C}\left(1-\frac{P}{CR}\right)$ (where R is the number of rows, C the number of columns, and P the number of pluses in the table), and the ratio to this of the sum of squared differences between observed and expected column counts is distributed as χ^2 with $C-1$ degrees of freedom, provided R is not too small and the proportion of plus signs is not too far from one-half (when R is small or the number of pluses and minuses extremely unequal, exact distributions may be computed without great difficulty). If the appropriate null hypothesis is that the observations are sequentially random, we know of no valid procedure except the laborious one of tabulating exact distributions, unless the procedure assuming randomly arranged differences turns out to be satisfactory also for randomly arranged observations.

The desirability of a solution to this problem is enhanced by the possibility of using the same criterion, signs of differences, to test the randomness of the separate series. On the other hand, the technique would be restricted in scope by the fact that in many time series problems neither the hypothesis of sequential randomness in the observations nor that of sequential randomness in the differences is tenable. For example, in testing the correlation between two economic series we should allow for the presence of cyclical characteristics which tend to duplicate, or to cancel, any chance coincidences. (In these cases, methods suggested earlier may be preferable—arranging the pairs in order according to one variate, and counting the number of signs of one kind, the number of reversals, or the number of phases of each duration in the other variate.)

IV

The generality and simplicity of tests based on signs of differences is due to the fact that they utilize only information that is inherent in every sample. When there exists relevant information other than that contained in the sample, as for example the knowledge that the observations are from a normal or approximately normal population, the tests are "inefficient" in the sense that they fail to utilize the supplementary information. Where no such additional knowledge exists, however, they are more effective than tests of narrower scope, in that they do extract some information from samples to which the less general tests cannot be applied at all. Furthermore, even when some supple-

mentary information is available, a test based on signs of differences may be quite efficient. For example, in certain types of data the signs of differences are more accurate than the magnitudes of either the observations or the differences.[11] Again, the problem studied may be such that direction of movement *per se* is significant and the magnitude of movements is irrelevant. A similar situation arises when it is desired to abstract from the effects of trend, for a test based on signs of differences is relatively insensitive to movements that are gradual in relation to the other movements in a series.

The greater the random movements in a series, the more likely are tests based on signs of differences to overlook systematic movements. This is true of any significance test, for all significance tests judge the systematic tendencies by comparison with random fluctuations. The weakness of tests based on signs of differences is, however, that the random variation ordinarily is reduced only slightly by increasing the amount of data; that is, the "averaging out" process is not very effective. It is possible to conceive of cases where increasing the length of a series would be highly effective (for example, in data having a trend rise of ten units per year, if the trend-adjusted data fall within a range of one unit), and also of cases where it would have absolutely no effect (for example, in data having a trend rise of one unit per year, if the trend-adjusted data never differ by less than ten units). The series on sweet potato acreage discussed previously lies toward the first, that on sweet potato production toward the second, of these extremes.

No general statement about the efficiency or inefficiency of tests based on signs of differences can be valid except in the particular context of a specific problem and body of data.

[11] With economic measures of the kind for which index numbers are ordinarily used, it may be certain that a change has been in a given direction (e.g., when all components of the index change in the same way), questionable how much the change has been (because of ambiguities in the weighting system), and meaningless to state an absolute standing. Similarly, the short-time directions of change of an index may be more meaningful than the long-time directions of change.

AN APPROACH TO THE MEASUREMENT OF FARM POPULATION PRESSURE IN WISCONSIN

By Thomas C. McCormick
University of Wisconsin

The concept of an optimum population is actually so basic in current population theory that without it the latter would have little meaning or motive. It is, however, still essentially a philosophical abstraction that is sometimes discounted or even denied. In spite of its important social consequences, the idea has never been the subject of serious quantitative testing and study in concrete situations. The present paper reports the results of a quantitative approach to one aspect of the problem recently made in Wisconsin, which may be of interest as a methodological venture.

From the agricultural point of view, an optimum population implies that the per capita income of a farm population is in part a function of the ratio of the size of the population to the amount of land farmed. Accordingly, if interfering factors can be controlled, variations in the acreage of farm land per capita of the farm population from county[1] to county should be accompanied by differences in the per capita farm income, which should range from a minimum to a maximum.

Only data published by the state and federal governments were available. Since as a rule such large-scale data can come only from government sources, this was regarded as one of the conditions of the problem.

The search for an index of the income of the farm population was able to discover nothing better than a detailed and careful estimate of gross farm income by counties made by the Wisconsin Department of Agriculture and Markets in 1931 (no later estimates have been prepared). These county estimates were deflated and reduced to a per capita basis of the rural farm populations[2] in 1930. Although the index is not without faults, it was found to be rather highly correlated[3] with other indexes of agricultural prosperity, such as the value of crops and livestock products per capita of the rural-farm population in 1930 ($r^2 = .81$), the percentage of farm homes with automobile, radio, and telephone ($r^2 = .62$), and the per capital value of farmers' dwellings

[1] The county was the smallest unit for which data were published. Data were lacking for a time series analysis.

[2] *Wisconsin Agriculture*, Bulletin No. 140, p. 45, 1932; and 1930 Census of the United States, *Population*, Vol. III, Pt. 2, p. 1328. The gross farm income is the estimated value of all products grown. This was divided by 146 and multiplied by 100 to give the deflated index. The index of prices paid by farmers for commodities bought in 1930 is 146, with 1910–14 as the base period. See *Wisconsin Agriculture*, *op. cit.*, p. 10, column 20.

[3] Computed from a test sample.

($r^2 = .64$). This index, because it represents the dependent variable, is represented by the symbol X_0, and will be referred to as "per capita income." The independent variable, the acreage of farm land per capita of the rural-farm population, was of course readily obtained for each county by dividing the total farm acreage by the rural-farm population[4] in 1930. It has the symbol X_1, and will be called the "land-man ratio." In studying the relationship between the per capita income (X_0) and the land-man ratio (X_1), the chief difficulty lay in the control of interfering factors. As usual in social research, this problem was only partially solved by holding constant farm land values per acre[5] (X_2) and type of farming[6] (X_3).

GENERAL METHOD AND RESULTS

The data and method of analysis are shown in Tables I through VI. Omitting the two highly urbanized counties of Milwaukee and Kenosha, the remaining sixty-nine counties in Wisconsin are divided into three broad types of farming areas, each area occupying two tables to itself. In Tables I, III, and V the counties in each area are classified first by mean value of farm land per acre (X_2), and second by land-man ratio (X_1). In any land-value class (row) in an area table (Tables I, III, and V), both type of farming and land value per acre are held roughly constant, so that variations in mean per capita income ($\overline{X}_0$) from one land-man ratio class (column) to another can probably be attributed to differences in the land-man ratio (X_1) and (unknown) associated factors. It can be seen by inspection that in every row (of the odd-numbered tables) the mean per capita income (X_0) increases as the land-man ratio (X_1) increases. The actual amount of this relationship, with land value (X_2) and type of farming (X_3) constant, is roughly measured by the squared partial correlation ratio, $\eta_{01.23}^2$, calculated for each row and averaged over the table. The $\eta_{01.23}^2$ values for Tables I, III, and V are .63, .68, and .49, respectively.

On the strength of the relationship just demonstrated between per capita income (X_0) and the land-man ratio (X_1), it is assumed that those counties that have the highest land-man ratio in a given row represent, from the point of view of per capita farm income, the most favorable adjustment of farm population to farm land in their class; and the mean land-man ratio of these "top" counties is taken as the criterion for adjusting upward the land-man ratios of the other counties

[4] 1930 Federal Census, *Agriculture*, Vol. II, Pt. 1, p. 736.

[5] 1930 Federal Census, *Agriculture*, Vol. II, Pt. 1, pp. 736, 748. Deflated by dividing value per acre by 115 and multiplying by 100. The index of farm real estate values in 1930 is 115, with 1910–14 as base period. See *Wisconsin Agriculture*, *op. cit.*, p. 10, column 22.

[6] Type of farming areas were taken with modifications from P. E. McNall and W. J. Roth, *Forces Affecting Wisconsin Agriculture, with Resulting Types of Farming*, Research Bulletin 131, Agricultural Experiment Station of the University of Wisconsin, November, 1935.

in the same class (row). The farm population required to raise the land-man ratio of a county to that of the criterion counties in the same row is found by dividing the farm acreage of the county in question by the mean land-man ratio of the criterion counties.[7] This is called the "adjusted" farm population. A county index of relative population pressure

TABLE I

TYPE OF FARMING AREA I, TWENTY-TWO WISCONSIN COUNTIES: METHOD OF ANALYSIS*

X_1	X_2	18–23	24–29	30–33	Totals
30–44	n	3	0	2	5
	$\bar{X}_3$	214.3		246.5	227.2
	ΣX_3	643		493	1136
	ΣX_3^2	143,605		121,549	265,154
	$(\Sigma X_3)^2/n$	137,816		121,525	258,099
45–59	n	8	4	1	13
	$\bar{X}_3$	214.9	303.0	328.0	250.7
	ΣX_3	1,719	1,212	328	3,259
	ΣX_3^2	376,641	369,882	107,584	854,107
	$(\Sigma X_3)^2/n$	369,370	367,236	107,584	817,006
60–84	n	4	0	0	4
	$\bar{X}_3$	252.8			252.8
	ΣX_3	1,011			1,011
	ΣX_3^2	256,677			256,677
	$(\Sigma X_3)^2/n$	255,530			255,530
Totals	n	15	4	3	22
	$\bar{X}_3$	224.9	303.0	273.7	245.7
	ΣX_3	3,373	1,212	821	5,406
	ΣX_3^2	776,923	369,882	229,133	1,375,938
	$(\Sigma X_3)^2/n_g$	758,475	367,236	224,680	1,350,391
	$\Sigma^r[(\Sigma X_3)^2/n_r]$				1,330,635
	$\Sigma^{rc}[(\Sigma X_3)^2/n_{rc}]$	762,716	367,236	229,109	1,359,061
	$(\Sigma X_3)^2/N$				1,328,402

* For meaning of symbols, see text.

NOTE: Tables I, III, and V would be somewhat more exact if the calculations had been weighted by the rural farm populations of the counties. In the case of estimates as rough as these, however, the extra labor required did not seem justified.

in agriculture, symbolized by the letter I, is then obtained[8] by dividing the adjusted farm population by the census rural farm population in 1930, multiplying by 100, and subtracting from 200. (See column I, Tables II, IV and VI.) Thus, an index of 120 means that the rural farm population of a county must be reduced by 20 per cent in order to produce a land-man ratio equal to that of the criterion counties in the same land-value class and type of farming area, which of course have an index of 100.

[7] Or, the adjusted farm population may be obtained by multiplying the observed rural farm population of the county by its index of population pressure, I, as described in the text, subtracted from 200 and expressed as a decimal.

[8] Or, the land-man ratio of the county may be divided by the criterion land-man ratio and the quotient subtracted from 2 and multiplied by 100.

TABLE II

TYPE OF FARMING AREA I: OBSERVED DATA AND ESTIMATES*

No.	County	X_0	X_1	X_2	P	$F.I.$	I	X_0'	P'	$F'.I.'$ $=X_0' \times P'$
1	Jefferson	276	23.8	44	13,604	3,759,863	126	302	10,067	3,040,234
2	Outagamie	190	18.8	44	17,843	3,398,904	141	231	10,527	2,431,737
3	Shawano	177	22.6	32	19,200	3,394,589	129	206	13,632	2,808,192
4	Columbia	250	32.1	33	13,876	3,469,247	100	250	13,876	3,469,247
5	Green Lake	243	33.0	38	6,248	1,516,849	100	243	6,248	1,516,849
6	Brown	184	17.9	49	16,534	3,035,274	134	311	10,912	3,393,633
7	Door	171	23.0	47	10,043	1,720,137	115	227	8,537	1,937,899
8	Fond du Lac	251	23.1	52	18,737	4,695,548	115	306	15,926	4,873,356
9	Kewaunee	188	21.6	51	9,601	1,807,808	120	264	7,681	2,027,784
10	Manitowoc	207	19.7	47	17,791	3,685,205	127	309	12,987	4,012,983
11	Sheboygan	235	18.7	56	15,980	3,748,288	131	351	11,026	3,870,126
12	Washington	257	20.2	51	12,903	3,312,192	125	352	9,677	3,406,304
13	Winnebago	226	21.1	50	12,004	2,712,671	122	309	9,363	2,893,167
14	Dane	273	24.7	54	28,763	7,856,644	100	273	28,763	7,856,644
15	Dodge	299	24.4	54	21,112	6,316,027	100	299	21,112	6,316,027
16	Rock	296	27.1	47	15,442	4,570 137	100	296	15,442	4,570,137
17	Walworth	344	27.5	50	11,380	3,919,726	100	344	11,380	3,919,726
18	Green	328	32.2	50	11,077	3,628,767	100	328	11,077	3,628,767
19	Calumet	229	19.5	61	9,864	2,259,041	100	229	9,864	2,259,041
20	Ozaukee	246	17.9	75	7,386	1,818,973	100	246	7,386	1,818,973
21	Racine	262	18.6	80	9,419	2,466,507	100	262	9,419	2,466,507
22	Waukesha	274	19.5	75	15,222	4,175,274	100	274	15,222	4,175,274
Totals					314,078	77,267,671			270,124	76,692,607
Means		246†	23.0	52			114†	284†		

* Meaning of symbols used as column headings:

X_0 is deflated gross farm income per capita of the rural farm population, 1931.

X_1 is acres of farm land per capita of the rural farm population, 1930, referred to as land-man ratio.

X_2 is the deflated value of farm land per acre, 1930.

P is the rural farm population, 1930.

$F.I.$ is the total deflated gross farm income, 1931.

I is the index of relative farm population pressure.

X_0' is the adjusted deflated gross farm income per capita of the rural farm population, 1931.

P' is the adjusted rural farm population, 1931.

$F'.I'.$ is the adjusted total deflated gross farm income, 1931.

† Weighted by the rural farm populations of the counties.

By this method absolute optima populations are avoided and relative optima, established by the counties themselves, are used that change as the value of farm land changes.

Next, a rough estimate is made of the probable effect of the increased land-man ratios on the farm income per capita in each adjusted county and on the total farm income in the type of farming area. The results show that while increasing the number of acres per capita always increases per capita income, in each of the three types of farming areas it reduces the total farm income because the returns from a fixed amount of farm land decrease as cultivation becomes less intensive (cf. the $F.I.$ and $F'.I.'$ columns of Tables II, IV, and VI). This appears again

in the fact that in Area II, for example, the coefficient of linear correlation between the land-man ratio and farm income *per acre* is $r = -.68$, or $r^2 = .46$. The operation of this principle is familiar in the thickly populated countries of Europe and Asia, where returns per acre are high but returns per worker are very low.

Finally, the investigation shows that in each type of farming area in Wisconsin relative over-population arises in greater degree from too many farm families than from too large farm families. There is a slight tendency for farm families to be larger in counties where the farms are smaller and incomes lower.

The steps in this analysis will be shown in some detail for type of farming Area I. In the cases of the other two areas only the end results will be given, except where certain points need explanation.

TYPE OF FARMING AREA I

Type of farming Area I, represented by Tables I and II, includes twenty-two counties located in the southern and eastern parts of the state. The dominant farm enterprise is dairying. The region is distinguished by the highest gross farm income per capita of the rural farm population, the weighted mean[9] being $246 (column X_0, Table II). The area also has the highest priced farm land, with a range of $32 to $80 and a mean of $52 per acre (column X_2, Table II). The mean land-man ratio, however, is only 23 acres of farm land per capita of the rural farm population, which is smaller than that of either of the other two areas (column X_1, Table II).

The relationship between per capita income (X_0) and land-man ratio (X_1) in this area, with type of farming (X_3) and value of land (X_2) constant, may be found from the formula

$$\eta^2_{01.23} = \frac{\sum^{r}\sum^{c}\frac{\left(\sum^{rc} X_0\right)^2}{n_{rc}} - \sum^{r}\frac{\left(\sum^{c} X_0\right)^2}{n_r}}{\sum^{r}\sum^{c} X_0^2 - \sum^{r}\frac{\left(\sum^{c} X_0\right)^2}{n_r}} \qquad (1)$$

where r refers to row, c to column, rc to cell, and n to frequency. In using this formula it is well to omit (from both numerator and denominator) any row that has less than two cells containing two or more X_0 entries each. Substituting from Table I,[10]

[9] Unless otherwise stated, all area means are unweighted averages of the county means. Weighting was usually found to make no practical difference in the results.

[10] It seemed advisable to avoid Bessel's correction and the correction for broad grouping. The assumptions of the latter make it invalid in this study, and since the two corrections tend to offset one another, it would be misleading to use the former alone. If both were made in the usual way, the values of η^2 would be materially raised here.

$$\eta^2_{01.23} = \frac{1{,}359{,}061 - 1{,}330{,}635}{1{,}375{,}938 - 1{,}330{,}635} = .63.$$

That is, roughly 63 per cent of the variance (standard deviation squared) in per capita income from county to county may be attributed to differences in the land-man ratio and associated factors, with the influence of differences in land values and type of farming removed.

It is also possible to estimate the extent to which per capita farm income reflects the combined influence of the two factors, land-man ratio, X_1, and land value per acre, X_2, with type of farming, X_3, held constant. For this purpose the following formula may be devised,

$$H^2_{0.12(3)} = 1 - \frac{\sum^r \sum^c X_0^2 - \sum^r \sum^c \frac{\left(\sum^{rc} X_0\right)^2}{n_{rc}}}{\sum^r \sum^c X_0^2 - \frac{\left(\sum^N X_0\right)^2}{N}} \qquad (2)$$

where N is the total frequency of the table. In making such an estimate, cells that contain less than two entries should be omitted entirely. The value here is

$$H^2_{0.12(3)} = 1 - \frac{1{,}375{,}938 - 1{,}359{,}061}{1{,}375{,}938 - 107{,}584 - \frac{(5406 - 328)^2}{21}} = .58.$$

But $H^2_{0.12(3)}$ should be greater than $\eta^2_{01.23}$ if the factors X_1 and X_2 are held actually constant in the cells. As the X_1 and X_2 categories in Table I become narrower the value of $H^2_{0.12(3)}$ increases. To correct $H^2_{0.12(3)}$ for broad categories we may find $\eta^2_{02.3}$ for the whole table

$$\eta^2_{02.3} = \frac{\sum^r \frac{\left(\sum^c X_0\right)^2}{n_r} - \frac{\left(\sum^N X_0\right)^2}{N}}{\sum^r \sum^c X_0^2 - \frac{\left(\sum^N X_0\right)^2}{N}} \qquad (3)$$

$$\eta^2_{02.3} = \frac{1{,}330{,}635 - 1{,}328{,}402}{1{,}375{,}938 - 1{,}328{,}402} = .05$$

and then use the formula

$$H^2_{0.12(3)} = \eta^2_{01.23}(1 - \eta^2_{02.3}) + \eta^2_{02.3} \tag{4}$$

$$H^2_{0.12(3)} = .63(.95) + .05 = .65.$$

We therefore assume that something like 65 per cent of the variance in per capita income between the counties is attributable to the combined influence of the land-man ratio and land values, leaving 35 per cent unexplained.

By a formula similar to (1),

$$\eta^2_{02.13} = \frac{\sum^{r}\sum^{c}\frac{\left(\sum^{rc} X_0\right)^2}{n_{rc}} - \sum^{c}\frac{\left(\sum^{r} X_0\right)^2}{n_c}}{\sum^{r}\sum^{c} X_0^2 - \sum^{c}\frac{\left(\sum^{r} X_0\right)^2}{n_c}} \tag{5}$$

$$\eta^2_{02.13} = \frac{1{,}359{,}061 - 1{,}350{,}391}{1{,}375{,}938 - 1{,}350{,}391} = .34.$$

A comparison of $\eta^2_{02.13} = .34$ with $\eta^2_{01.23} = .63$ shows that the relationship between the land-man ratio and per capita income is much closer than that between the per capita income and the value of farm land per acre.

In Table I it is seen that five counties have farm land with mean values ranging between $30 and $45 an acre. Three of the counties, Shawano, Outagamie, and Jefferson, have a mean land-man ratio of 21.7 acres and a mean per capita income of $214. The other two counties, Green Lake and Columbia, have a mean land-man ratio of 32.6 acres and a per capita income of $247. If, in view of our correlation analysis above, we attribute the difference of about $32 in per capita income between these two groups of counties to the difference in their land-man ratios (and unknown associated factors), we may estimate that each increase of one acre in land-man ratio is accompanied by a mean gain of roughly $3.13 in per capita income $\left(\frac{246.5-214.3}{32.0-21.7}\right)$. To raise the mean land-man ratio of the three "low" counties to 32 acres, we divide 32 into the land-man ratio of each of these counties, i.e., into 22.6 in the case of Shawano, into 18.8 in the case of Outagamie, and into 23.8 in the case of Jefferson, and subtract the resulting percentage from 200 in each case. We then obtain as the county indexes of relative farm population pressure in this land-value class: Jefferson 126, Outagamie 141, Sha-

wano 129 (column I, Table II). That is, the rural farm population of Jefferson county would have to be reduced 26 per cent in order to provide 32 acres of farm land for the support of each individual member. Columbia and Green Lake, the criterion counties in this group, each has a population pressure index of 100, and their farm populations are left unchanged. When the farm populations of Jefferson, Shawano, and Outagamie counties are reduced in this way, the adjusted farm population for each of the three counties may be found by multiplying its 1930 rural farm population by the difference between 200 and the county's index of population pressure, expressed as a decimal. For example, for Jefferson county, with a rural farm population of 13,604 in 1930, the adjusted population is $13{,}604 \times (2.00 - 1.26) = 10{,}067$. For the other two counties the 1930 rural farm populations and their adjusted values are: Outagamie, 17,843 reduced to 10,527; Shawano, 19,200 reduced to 13,632 (column P', Table II).

If such adjustments in the land-man ratios of Jefferson, Outagamie, and Shawano counties were made, we infer that each county would probably have its per capita farm income increased by an amount equal roughly to the increase in its land-man ratio multiplied by the figure \$3.13 obtained above. Thus, in the case of Jefferson county, the land-man ratio is increased from 23.8 acres to 32 acres, or 8.2 acres, so that the gain in its per capita income is $8.2 \times 3.13 = \$25.67$. Since the observed per capita income of Jefferson is \$276, its adjusted per capita income is $276 + 25.67 = \$301.67$, or about \$302. Similarly, in Outagamie county the per capita income is raised from \$190 to \$231, and in Shawano county from \$177 to \$206 (column X_0', Table II).

By repeating these simple calculations for each land-value class in Table I, Table II (to the right of the upright double ruling) was worked out for the twenty-two counties in Area I. The land-man ratio of each of the counties numbered 6 through 13 in Table II was raised to 27 acres, and the addition of each acre was estimated to increase the per capita income by \$14.00.

The weighted index of farm population pressure for the whole of Area I is 114 (Table II). It is found by dividing the total observed rural farm population (P) into the total adjusted farm population (P'), and subtracting from 200, or $200 - \frac{270{,}124}{314{,}078} = 114$ (columns P and P' in Table II). That is, a reduction of 14 per cent in the rural farm population of this prosperous dairy section in southern and eastern Wisconsin is needed to raise the land-man ratios in all counties to a par with the land-man ratios observed in the counties with the highest land-man ratios in their respective land-value classes. A population of about 44,000 (i.e.,

314,078×.14) would be released from the farms, chiefly through migration to the cities, partly through birth control. As a result of such a reduction in the rural farm population, it is likely that the total farm income for the area would be decreased from $77,267,700 to $76,692,600 (columns *F.I.* and *F.'I.'*, Table II), a loss of $575,000, or a little less than one per cent. The weighted per capita income of the remaining farm population, however, would be raised about $38, or 15 per cent, as a result of an increase from $246 $\left(\text{i.e., } \frac{77{,}267{,}671}{314{,}078}\right)$ to $284 $\left(\text{i.e., } \frac{76{,}692{,}607}{270{,}124}\right)$ (columns *P*, *F.I.*, *P'*, and *F.'I.'*, Table II).

TYPE OF FARMING AREA II

Type of farming Area II (Tables III and IV) lies in southern and western Wisconsin, and also contains twenty-two counties. It is a region of dairy, livestock, and mixed farming. Farm land ranges in mean value from $28 to $46 an acre, with a mean for the area of $35 an acre. The weighted gross farm income per capita of the rural farm population is $199. It is somewhat poorer than Area I but more prosperous than Area III. The land-man ratio averages 28.8 acres per capita of the rural farm population, compared with 23 acres in Area I. The values $\eta^2_{02.13} = .33$; $\eta^2_{01.23} = .68$; and $H^2_{0.12(3)} = .71$ closely resemble the corresponding values found for Area I, and again indicate a strong relationship between per capita farm income and the land-man ratio. The estimates for

TABLE III

TYPE OF FARMING AREA II, TWENTY-TWO WISCONSIN COUNTIES: METHOD OF ANALYSIS*

X_1	X_1	20–25	26–31	32–37	38–41	Totals
25–39	n	7	7	1	2	17
	$\bar{X}_0$	181.1	203.9	196.0	245.0	198.9
	ΣX_0	1,268	1,427	196	490	3,381
	ΣX_0^2	231,394	292,999	38,416	120,212	683,021
	$(\Sigma X_0)^2/n$	229,689	290,904	38,416	120,050	672,421
40–49	n	1	2	2	0	5
	$\bar{X}_0$	181.0	224.0	245.5		224.0
	ΣX_0	181	448	491		1,120
	ΣX_0^2	32,761	101,152	120,292		254,205
	$(\Sigma X_0)^2/n$	32,761	100,352	120,541		250,880
Totals	n	8	9	3	2	22
	$\bar{X}_0$	181.1	208.3	229.0	245.0	205.0
	ΣX_0	1,449	1,875	687	490	4,501
	ΣX_0^2	264,155	394,151	158,708	120,212	937,226
	$(\Sigma X_0)^2/n_c$	262,450	390,652	157,323	120,050	930,448
	$\Sigma^r[(\Sigma X_0)^2/n_r]$					923,301
	$\Sigma^{rc}[(\Sigma X_0)^2/n_{rc}]$	262,450	391,256	158,957	120,050	932,713
	$(\Sigma X_0)^2/N$					920,864

* For meaning of symbols, see text.

TABLE IV

TYPE OF FARMING AREA II: OBSERVED DATA AND ESTIMATES*

No.	County	X_0	X_1	X_2	P	$F.I.$	I	X_0'	P'	$F.I.' = X_0' \times P'$
1	Barron	189	21.6	36	21,916	4,012,329	146	258	11,446	2,953,068
2	Chippewa	183	25.8	33	18,131	3,326,164	135	236	11,785	2,781,260
3	Clark	187	20.5	37	24,533	4,588 151	149	260	12,512	3,253,120
4	Marathon	162	20.7	35	34,455	5,568,151	148	234	17,917	4,192,578
5	Pierce	211	25.5	37	14,432	2,827,877	136	265	9,236	2,447,540
6	Polk	167	23.5	34	18,953	3,173,630	141	229	11,182	2,560,678
7	Wood	169	22.3	37	13,725	2,319,452	144	213	7,686	1,637,118
8	Dunn	204	28.8	34	16,899	3,441,986	128	246	12,167	2,993,082
9	Eau Claire	181	28.5	31	10,352	1,878,082	129	224	7,350	1,646,400
10	La Crosse	236	30.8	32	8,877	2,091,438	123	271	6,835	1,852,285
11	Monroe	204	29.4	32	15,833	3,223,356	122	244	12,350	3,013,400
12	Pepin	183	28.8	31	4,677	853,836	128	225	3,367	757,575
13	St. Croix	214	28.5	30	15,018	3,208,562	129	257	10,663	2,740,391
14	Trempeleau	205	29.4	31	15,100	3,102,466	126	245	11,174	2,737,630
15	Crawford	196	35.2	28	9,323	1,824,384	112	214	8,204	1,755,656
16	Buffalo	254	40.3	28	10,134	2,572,740	100	254	10,134	2,572,740
17	Iowa	236	39.5	39	11,557	2,725,274	100	236	11,557	2,725,274
18	Vernon	181	25.5	42	19,193	3,476,233	125	229	14,395	3,296,455
19	Richland	204	29.5	40	12,146	2,479,452	113	229	10,567	2,419,843
20	Sauk	244	30.2	40	16,232	3,953,082	111	265	14,446	3,828,190
21	Grant	234	34.6	44	19,594	4,578,699	100	234	19,594	4,578,699
22	Lafayette	257	34.1	46	10,995	2,843,425	100	257	10,995	2,843,425
Totals					341,355	68,068,769			245,562	59,586,407
Means		199†	28.8	35			128†	243†		

* For meaning of symbols, see Table II.

† Weighted by the rural farm populations of the counties.

Area II are collected in Table IV. A land-man ratio of 40 acres was taken as the criterion for counties numbered 1 through 17; while for counties numbered 18 through 22 the criterion was 34 acres. In the first case the addition of each acre to the land-man ratio increased the per capita income by about $3.75, in the second case by about $5.62.

Area II shows a weighted index of farm population pressure of 128, compared with 114 for Area I. To reduce this index to 100 would require the removal of some 95,800 people from the farms of the twenty-two counties. Such a reduction in the farm population of the area would be accompanied by a loss in total gross farm income estimated at 12 or 13 per cent ($8,482,000). On the other hand, the weighted per capita income would be increased 22 per cent, from $199 to $243.

TYPE OF FARMING AREA III

The third and last type of farming area, covered by Tables V and VI, includes twenty-five counties. Nineteen of these counties fall in the northern part of the state, six in the central part. The area includes the northwestern sandy dairy-potato and dairy-hay regions, the central

TABLE V

TYPE OF FARMING AREA III, TWENTY-FIVE WISCONSIN COUNTIES: METHOD OF ANALYSIS*

X_1	X_1	14–19	20–25	26–31	32–37	38–47	Totals
10–24	n	3	5	3	1	2	14
	$\bar{X}_3$	83.0	118.4	144.3	170.0	153.5	125.1
	ΣX_3	249	592	433	170	307	1,751
	ΣX_3^2	20,981	71,110	64,769	28,900	48,220	233,980
	$(\Sigma X_3)^2/n$	20,667	70,093	62,496	28,900	47,125	219,100
25–39	n	3	7	1	0	0	11
	$\bar{X}_3$	116.0	136.1	165.0			133.3
	ΣX_3	348	953	165			1,466
	ΣX_3^2	43,322	135,863	27,225			206,410
	$(\Sigma X_3)^2/n$	40,368	129,744	27,225			195,459
Totals	n	6	12	4	1	2	25
	$\bar{X}_3$	99.5	128.8	149.5	170.0	153.5	128.7
	ΣX_3	597	1,545	598	170	307	3,217
	ΣX_3^2	64,303	206,973	91,994	28,900	48,220	440,390
	$(\Sigma X_3)^2/n_c$	59,400	199,068	89,400	28,900	47,125	423,893
	$\Sigma^r[(\Sigma X_3)^2/n_r]$						414,559
	$\Sigma^{rc}[(\Sigma X_3)^2/n_{rc}]$	61,035	199,837	89,721	28,900	47,125	426,618
	$(\Sigma X_3)^2/N$						413,964

* For meaning of symbols, see text.

TABLE VI

TYPE OF FARMING AREA III: OBSERVED DATA AND ESTIMATES*

No.	County	X_3	X_1	X_2	P	$F.I.$	I	X_3'	P'	$F.'I.'$ $=X_3' \times P'$
1	Forest	84	19.6	22	3,116	263,219	156	125	1,371	171,375
2	Iron	70	14.7	22	2,995	210,411	167	119	988	117,572
3	Price	95	19.1	19	9,561	910,890	158	137	4,016	550,192
4	Ashland	123	21.6	23	5,086	624,109	152	161	2,441	393,001
5	Bayfield	126	23.2	22	8,338	1,053,014	148	161	4,336	698,096
6	Florence	126	24.3	20	1,664	209,315	146	160	899	143,840
7	Marinette	127	21.7	22	12,012	1,522,877	152	165	5,766	951,390
8	Sawyer	90	22.9	23	4,467	403,973	149	126	2,278	287,028
9	Burnett	134	27.4	21	7,574	1,014,041	139	163	4,620	753,060
10	Jackson	182	31.1	23	11,244	2,042,740	131	205	7,758	1,590,390
11	Washburn	117	28.7	21	5,860	683,151	136	143	3,750	536,250
12	Waushara	170	34.4	21	9,527	1,618,630	124	187	7,241	1,354,067
13	Adams	130	46.7	13	5,524	718,699	100	130	5,524	718,699
14	Marquette	177	43.6	18	5,781	1,020,479	100	177	5,781	1,020,479
15	Langlade	143	18.9	35	9,284	1,320,479	139	188	5,663	1,064,644
16	Taylor	133	19.7	28	12,519	1,666,849	136	175	8,012	1,402,100
17	Vilas	72	16.2	27	2,298	166,370	148	127	1,195	151,765
18	Douglas	111	21.6	25	7,870	873,836	130	146	5,509	804,314
19	Lincoln	134	23.3	29	8,195	1,101,781	125	163	6,146	1,001,798
20	Oconto	141	20.6	31	14,812	2,087,534	134	180	9,776	1,759,680
21	Oneida	104	25.2	27	3,936	407,945	119	126	3,188	401,688
22	Portage	127	25.8	28	16,256	2,063,904	117	146	13,492	1,969,832
23	Rusk	134	20.3	26	9,844	1,315,959	135	174	6,399	1,113,426
24	Waupaca	202	25.1	35	16,297	3,299,658	119	224	13,201	2,957,024
25	Juneau	165	31.4	26	9,690	1,597,329	100	165	9,690	1,597,329
Totals					203,750	28,197,192			139,040	23,509,039
Means		138†	25.1	24			132†	169†		

* For meaning of symbols, see Table II.

† Weighted by the rural farm populations of the counties.

sandy general farming region, and the central sandy-loam dairy-potato region. Farm land values are much lower than in either of the other two areas, ranging from $10 to $39 per acre, with a mean of $24 per acre. The weighted per capita farm income, only $138, is also the smallest among the three areas. The mean land-man ratio of 25 acres falls between the land-man ratios of Areas II and I. The values of the correlation ratios are somewhat less than those for the other two areas: $\eta^2_{01.23} = .49$; $\eta^2_{02.13} = .19$; $H^2_{0.12(3)} = .50$. However, there is again a substantial relationship between per capita farm income and land-man ratio. The criteria used were: among counties numbered from 1 through 14 in Table VI, a land-man ratio of 45 acres, with an increase in per capita income of $1.62 for each acre added to the land-man ratio; among counties numbered from 15 through 25, a land-man ratio of 31 acres, with an increase of $3.73 in per capita income for each acre added to the land man ratio.

The weighted index of farm population pressure for Area III is 132, which is the highest among the three areas. Some 64,700 persons would have to be withdrawn from the rural farm population in order that all counties in the area might attain the land-man ratios of the criterion counties. Such a population adjustment would result in an estimated reduction of about 17 per cent ($4,688,000) in the total gross farm income, but would raise the weighted mean farm income per capita 22 per cent, from $138 to $169.

CHANGES IN COUNTY FARM POPULATIONS, 1930–1940, IN RELATION TO 1931 INDEXES OF POPULATION PRESSURE

Between the censuses of 1930 and 1940 the rural-farm populations of the counties with indexes of population pressure greater than 100 in 1931 decreased about 3 per cent on the average in Area I and also in Area II; whereas the farm population of the criterion counties with indexes of 100 increased about 3 per cent in Area I and remained practically stationary in Area II. In type of farming Area III the farm population of the counties with indexes of population pressure above 100 increased nearly 5 per cent over the same period, while that of the criterion counties increased only about one per cent.

These figures suggest that in Area I adjustments in the size of the farm population have recently taken place that are tending to equalize population pressure on agricultural resources throughout the area at an average level above that of the criterion counties but below that of the counties with population indexes in excess of 100 in 1931. In type of farming Area II the slight recession of farm population in the criterion counties and the more marked decline in the other counties is also work-

ing toward equalization of pressure, but at a level below that of the criterion counties in 1931. By contrast, in Area III the inequalities of 1931 between the criterion counties and the more crowded counties have been increasing. This is further evidence that the major problem of population pressure on agricultural resources is to be found in the northern and central parts of the state, particularly in the so-called "Cut-Over" region.

SIZE OF FARM VS. SIZE OF FARM FAMILY IN RELATION TO INDEX OF FARM POPULATION PRESSURE

When the indexes of farm population pressure are linearly correlated with the percentage of young people under twenty years of age[11] in the rural-farm population in 1930 by counties in each type of farming area, and the three coefficients so obtained are averaged, we find $\bar{r}^2 = .25$. Similarly, between the indexes of population pressure and size of farm family[12] the mean $\bar{r}^2 = .14$; and between the indexes of population pressure and mean size of farm[13] the mean $\bar{r} = -.91$, or $\bar{r}^2 = .83$. If we represent the indexes of population pressure by X_4, the size of farm by X_5, and the size of farm family by X_6, and obtain coefficients of linear partial correlation, we find $\bar{r}_{45.6} = -.86$, or $\bar{r}^2_{45.6} = .74$; and $\bar{r}_{46.5} = .49$, or $\bar{r}^2_{46.5} = .24$. Finally, the simple linear correlation between size of farm and size of farm family is mean $\bar{r} = -.20$, or $\bar{r}^2 = .04$. The interpretation of these coefficients has been given above under "General Method and Results."

[11] 1930 Federal Census, *Population*, Vol. III, Pt. 2, p. 1307.
[12] Ibid., *Agriculture*, Vol. II, Pt. 1, p. 736.
[13] Ibid., *Population*, Vol. VI, *Families*, p. 1468.

NOTE: It will be noticed that all of the data used in this study were taken from the 1930 Federal Census, except the gross farm income by counties, which was furnished by the Wisconsin Department of Agriculture and Markets. Unfortunately, reliable estimates of farm income by counties are rarely available in any state, including Wisconsin. A fairly good substitute, however, may now be obtained from the Federal Census. This is the value of all farm products sold, traded, or used by farm households (*Agriculture*, Third Series, County Table XVII, 1940). The squared coefficient of correlation between these values of farm products and the gross farm incomes (X_3) used in this study, both on a per capita basis, is $r^2 = .81$, as already mentioned.

A study dependent on county data drawn solely from the Federal Census and publications of a state department of agriculture, of course, should not be used alone to recommend or guide any actual adjustments of the farm population to the land resources of a county or state. When such adjustments are made, as in the case of the resettlement projects of the Federal Government in northern Wisconsin, such studies as this should be supplemented by intensive economic, farm management, and sociological studies conducted in the field.

THE PRODUCTION FUNCTION FOR CANADIAN MANUFACTURES*

By Patricia Daly
University of Chicago
and Paul H. Douglas
Captain, U.S.M.C.R., and University of Chicago

Readers of this Journal may perhaps be somewhat acquainted with a series of articles by one of the present authors and various associates which on the basis of (1) time-series and (2) cross-section studies of labor, capital, and product have sought to discover by use of the production formula, $P=bL^kC^{1-k}$, the effects upon the quantity of manufacturing product caused by changes in the quantities of labor and capital. Thus far the economies studied have been the manufacturing industries of the United States[1] and Australia.[2] The present study gives empirical values to the parameters of the above function for the Canadian manufacturing industries. Four years have been chosen for the cross-section analysis,[3] namely, 1923, 1927, 1935, 1937, care being taken to avoid depression years and to select years which were relatively "normal."

In each of these given years, each manufacturing industry was treated as an observation with its respective quantities of Labor (L), Capital (C), and Product (P). The definitions given to these terms were as follows:

1. Labor (L) = total number employed, i.e., wage-earners and salaried workers.
2. Capital (C) = total capital, including both (a) fixed capital in plant, tools, machinery, etc., and (b) working capital including materials, goods in process, etc.

* The authors wish to thank their friends and colleagues, Miss Grace T. Gunn and Mr. H. Gregg Lewis for numerous helpful suggestions. We also wish to acknowledge aid given, in computing by E. Olson, B. Nimer, and W. L. Slayton, and in drawing of the charts by Mr. Y. K. Wong.

[1] Paul H. Douglas, *The Theory of Wages*, pp. 113–167; M. Bronfenbrenner and Paul H. Douglas, "Cross-Section Studies in the Cobb-Douglas Function," *Journal of Political Economy*, XLVII (December, 1939), pp. 761–785; Grace T. Gunn and Paul H. Douglas, "The Production Function for American Manufacturing in 1919," *American Economic Review*, XXXI (March, 1941), pp. 67–80; and "The Production Function for American Manufacturing for 1914," *Journal of Political Economy*, L (August, 1942), pp. 595–602; Patricia Daly, Ernest Olson, and Paul H. Douglas, "The Production Function for Manufacturing in the United States, 1904," *Journal of Political Economy* (forthcoming).

[2] Grace T. Gunn and Paul H. Douglas, "Further Measurements of Marginal Productivity," *Quarterly Journal of Economics*, LV (May, 1940), pp. 399–428; and "The Production Function for Australian Manufacturing," *ibid.*, LVI (November, 1941), pp. 108–129; Paul H. Douglas, *The Theory of Wages*, pp. 167–172; Paul H. Douglas and Marjorie L. Handsaker, "The Theory of Production as Derived from Australian Data," *Quarterly Journal of Economics*, LII (November, 1937, and March, 1938) pp. 1–36, 215–54.

[3] For the sources, see the Canada Year Book for 1925, 1930, 1938, and 1940.

3. Product (P) = gross sales value of product minus cost of materials used, fuel, electricity, etc.

The equation $P = bL^kC^j$ can be rewritten as a linear one by using the logarithmic values of the variables. The equation then becomes

$$\log P = \log b + k \log L + j \log C.$$

Fitting this equation by the method of least squares, the values in Table I are obtained.

It will be noticed:

1. That the sum of the exponents for labor and capital (i.e., $k+j$) closely approximate unity or 1.0. This is what has been found to be approximately true in virtually all of the other studies and strengthens the conclusion that the production function is homogeneous and of the first degree. This corroborates the deductive conclusion drawn that under competition production is carried to the point of lowest cost on the average total unit cost curve which is the point of constant costs.

2. That the exponents of both labor (k) and capital (j) are relatively steady. The span in values of k was from .43 to .50 which was less than the range of one standard error from the intermediate values. The span in values of j was from .48 to .58.

3. The values of k are somewhat less and the values of j somewhat more than those for the United States found for four particular years of the period 1904–1919 which were as follows:

Year	k	j
1904	.65	.31
1909	.74	.32
1914	.61	.36
1919	.76	.25

This is not surprising. We have never claimed that the exponents of the factors should be the same for all economies and periods. Indeed in *The Theory of Wages*, it was explicitly stated that one might expect such variations.[4] In the United States where capital per worker is appreciably greater we might expect capital to have a lower and labor a higher exponent than in Canada.

It is also interesting to note that the exponents of labor and capital for New Zealand as shown by a time study by Dr. Max Brown seem to approximate those of Canada.[5]

[4] See, Paul H. Douglas, *The Theory of Wages*, p. 203.

[5] The results of this study are given by Colin Clark in his book, *Conditions of Economic Progress*, p. 384. The material is contained in an unpublished thesis at Cambridge University, which through the courtesy of Dr. Brown we have had the privilege of reading.

We would conclude that the elasticity of the marginal productivity curve for labor in Canadian manufacturing for the four years studied was probably slightly under 2.0 and that for capital slightly over 2.0 (i.e., under competitive conditions the elasticity of this curve is equal to the reciprocal of the exponent of the other factor).

A test of the accuracy of the production function is the following: (1) compute for each industry from its actual quantities of Labor (L) and Capital (C) what the Product should be according to the formula. We shall call this P'; then (2) compare the actual product (P) with this computed product (P') and see how great are the relative differences

TABLE I

VALUES OF EXPONENTS OF LABOR AND CAPITAL IN THE PRODUCTION FUNCTION FOR CANADIAN MANUFACTURES*

Year	No. of Industries	k	σ_k	j	σ_j	b	$k+j$
1923	167	.48	.038	.48	.035	48.5	.96
1927	163	.46	.039	.52	.036	33.0	.98
1935	165	.50	.044	.52	.040	22.2	1.02
1937	164	.43	.041	.58	.036	15.4	1.01
Average for four years		.468		.525			.992

* If we use fixed capital only for C, the exponents are changed to the following:

Year	k	j	b	$k+j$
1923	.54	.44	82.0	.98
1927	.49	.49	58.9	.98
1935	.53	.49	33.5	1.02
1937	.48	.53	32.9	1.01
Average	.510	.488		.998

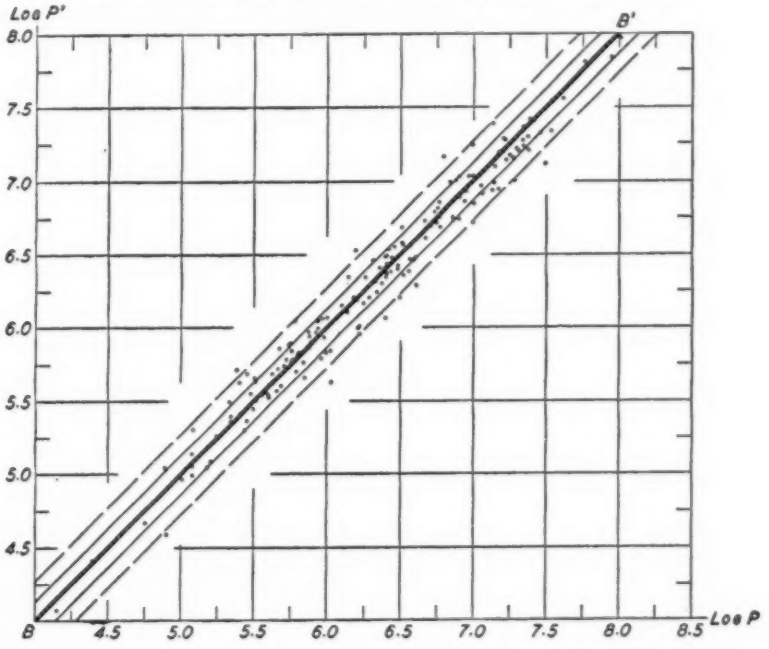
of the former from the latter and whether they can be explained by random errors of measurement or by other random causes operating upon the value product of specific industries.

This has been done for the years 1923, 1927, and 1935; the results being shown graphically in Charts I to III inclusive. The logs of the actual product or P are shown on the horizontal axis or abscissa; the logs of the computed product or P' upon the vertical axis or ordinate. Were there to be perfect correspondence between the two, each observation would fall on the line BB'. The differences between the actual product (P) and those which we would expect under the formula (P') are measured by the horizontal deviations of the observations from the line BB'. To measure these more accurately two sets of bands have been laid out on the charts. The first is at a distance of one standard error of estimate to the right and to the left of the line BB' while the second set of bands is at a distance of from one to two standard errors of estimate from the line of expected relationship.

It will be noticed that the observations cluster closely around the line BB' and that the scatter is not great. This is further evidence that the production formula furnishes a good description of the product in these manufacturing industries. In Table II there is shown the percentage of observations deviating from the line of perfect correspondence by less than one and two, and by more than two standard errors of estimate of P'. Chart IV is a graph of the results for 1937.

CHART I

THEORETICAL AND OBSERVED VALUES OF LOG P FOR 1923



We should not expect that there would be a precise agreement in each industry between P and P'. The modern theory of causation does not require any such one to one relationship. It is enough that given causal forces express themselves in "norms" of results or tendencies accompanied by a more or less symmetrical range of variations resulting from errors of sampling, errors of measurement, and sheer chance. That is one of the implications of the theory of probability. If the actual results do approximate the norms fairly closely and the variations are rather symmetrically distributed, then credence in the formula as both descriptive and predictive is greatly strengthened.

In the present case we would not expect deviations due to sampling

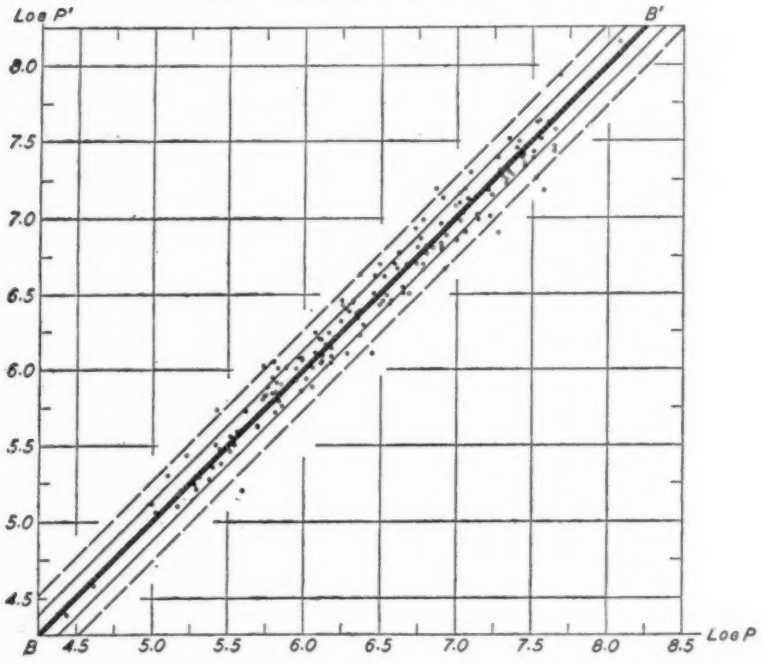
errors because of the fact that the quantities of L, C, and P are all derived from an all inclusive census and not from any sample. Some differences between P and P' would however inevitably result from:

(1) Errors of measurement.

(2) Shifts over time in the demand schedules for the end products of the industries causing some to move to the right and others to the left

CHART II

THEORETICAL AND OBSERVED VALUES OF LOG P FOR 1927



and hence affecting the money values of the products. Here it should be noted that under the quantity theory of money, if the supply of money and bank credit is constant, the average elasticity of demand is equal to unity. So is its reciprocal, the average flexibility of price. The variations in P caused by shifts of the demand curves will under these conditions tend to be mutually compensatory and hence will operate in a non-biased manner. The deviations from the norm caused by such shiftings will have, therefore, some tendency towards symmetry.

(3) Differences in the technical effectiveness of labor and capital in specific industries. We have never maintained that labor and capital must have the same exponents in each and every industry, but have

TABLE II

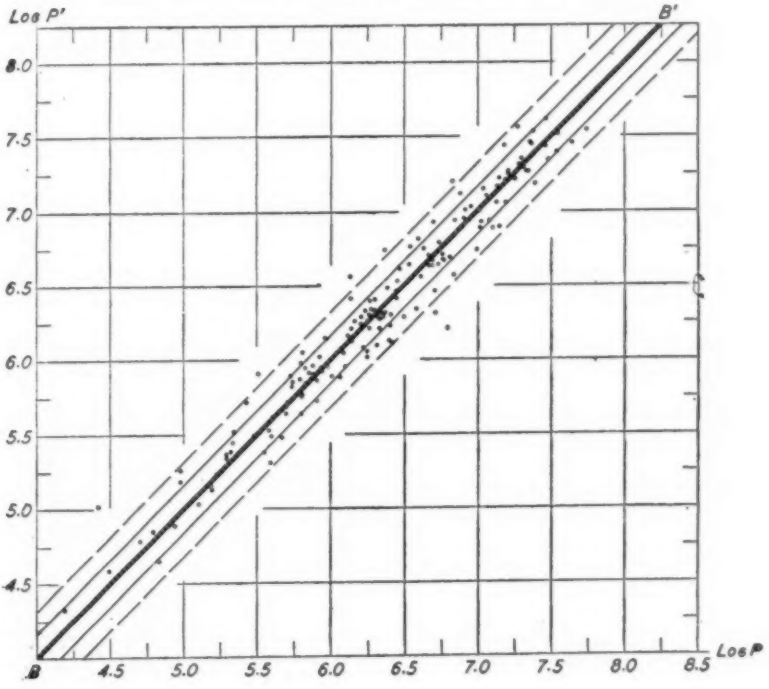
DEVIATIONS OF ACTUAL PRODUCT (P) IN CANADIAN MANUFACTURING FROM PRODUCT (P') COMPUTED UNDER THE FORMULA FOR THE YEARS 1923, 1927, 1935, AND 1937

Year	Percentage of observations of given deviations from P'		
	Within $\pm 1S$	Within $\pm 2S$	Outside $\pm 2S$
	Per cent	Per cent	Per cent
1923	72.46	92.81	7.19
1927	74.23	95.09	4.91
1935	72.17	95.76	4.24
1937	76.83	93.90	6.10

instead always pointed out these may well differ. The empirical exponents which we find are instead merely norms or averages of these varying exponents. It would be expected, therefore, that the actual products from industry to industry would not jibe precisely with the computed products based on the use of these "average" exponents of labor and capital. But, since the exponents for manufacturing industry as a whole are the norms or averages for the specific industries, the

CHART III

THEORETICAL AND OBSERVED VALUES OF LOG P FOR 1935

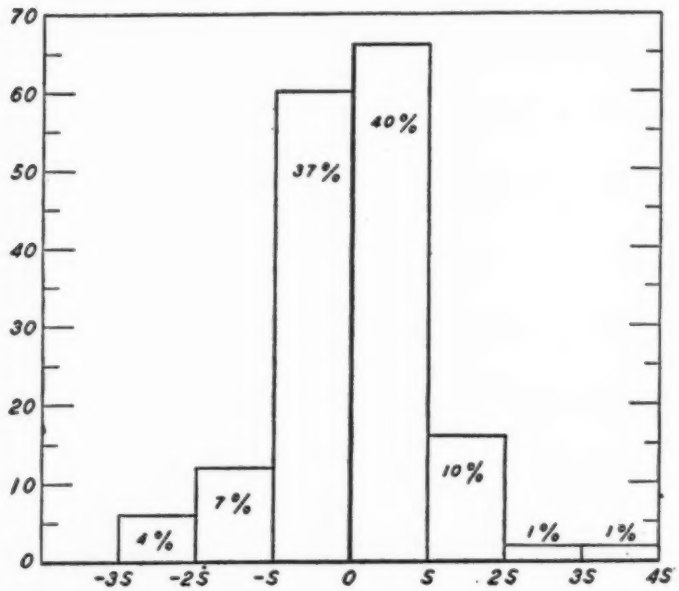


variations of these actual exponents of labor and capital in the individual industries would again be more or less symmetrically distributed and would not be biased.

(4) A fourth factor in causing differences between actual and computed P would be forces affecting production other than the quantities of labor and capital.

CHART IV

FREQUENCY DISTRIBUTION OF DEVIATIONS OF P FROM P' FOR 1937



(5) Finally, sheer chance itself may also be expected to play a part in causing individual deviations from the norms.

But, and this is crucially important, to the degree that these other causes operate at random and not in a biased manner, then the values of the exponents of labor and capital in the production function are the main determinants of the end result, namely, the value of the product. Under these conditions, we would expect that in two-thirds of the cases, the actual product would not deviate from the theoretical product by more than one standard error and that in 95 per cent of the cases, these deviations would not be more than two standard errors.

In practice we find that about 74 per cent of the observations fall within one standard error of the theoretical production values, and that in approximately 94 per cent of the cases, the deviation is less

than two standard errors. The relative differences between the actual and computed values of P are therefore at least not more, and in part actually less than would be expected under the theory of normal probability. Credence in the developed production function as a description of the relative effects of labor and capital upon product is thereby strengthened. The variations from it are normal.

A study of the relatively small number of industries where the deviations of actual from theoretical products were more than two standard errors of estimate shows that in the case of charcoal and coke, the actual product fell short of the theoretical by this amount in no less than three out of the four years; so too did clay products, linen goods, agricultural implements, and petroleum products in two out of the four years. On the other hand the actual values of abrasive products exceeded the theoretical values by more than two standard errors in three out of the four years. Also the actual values of cigars and cigarettes exceeded the theoretical values by more than two standard errors in two out of the four years. In view of the high degree of monopoly prevailing in this industry this is not surprising. Similarly in the case of breakfast foods, the money values of the actual products also exceeded the theoretical values by more than two standard errors in two of the years. In view of the importance of trade-marked brands creating a state of highly imperfect competition this again is what should be expected. Another industry where in two of the years the actual products were more than two standard errors above the theoretical values was the liquor industry.[6]

Some of the industries for which in one of the four years the actual value of product differed from the theoretical by more than two standard errors were toys, chopping mills, coal tar, fly paper, signs, paper patterns, bicycles and motorcycles, fur dressing and dyeing, maple syrup and sugar, fertilizers, musical instruments, fountain pens, nonferrous metals, sand lime brick, and wood distillation. On the whole a study of the deviations gives added proof of the general reliability of the formula and of its values for the four years studied.

A further test of the function can be made by comparing the actual share of the value product received by labor with that which we would expect to be received with the production formula. Under conditions of perfect competition one would expect labor to receive as its proportionate share of the total product the proportion represented by its own exponent k or more properly by the ratio $\frac{k}{k+j}$; $\frac{k}{k+j}$ being a more proper

[6] Distilled liquors in one year and malt liquors in another.

representation of labor's share of product under conditions of increasing or decreasing returns where $k+j$ is greater than or less than one.

In Table III we make this comparison, the share of labor being rep-

TABLE III

Year	W/P	k	$\frac{k}{k+j}$
1923	.50	.48	.50
1927	.48	.46	.47
1935	.40	.50	.49
1937	.52	.43	.43
Average	.475	.468	.472

resented by W/P or the ratio of total wages and salaries to the net value of the product. Taking the four years as a whole there was an almost precise correspondence between the share which labor actually received and that which we would expect it to receive. In the first two years, this agreement was very close, while in the last two years there were compensating differences. In 1935, the share which labor received was from 9 to 10 per cent less than one would expect from the formula while in 1937 it was 9 per cent more.

Taken in the large, the results of our Canadian studies seem to strengthen the case for the use of the production function.

THE COMPOSITION OF INCOME AND OWNERSHIP OF CAPITAL BY INCOME CLASSES IN THE UNITED STATES IN 1936

By Rufus S. Tucker
General Motors Corporation

Much current controversy revolves around the point of the distribution of income and wealth, and the alleged necessity of changing it, by taxation or other means. For that reason any statistical information concerning their distribution, if it is fairly close to accuracy, should be helpful as a substitute for prejudices and assumptions. A good deal of information relating to the source of income in the various income classes can be gleaned from the recently published special studies of income tax returns for 1936, filed in accordance with the income tax laws of the United States and of the states of Wisconsin and Delaware. The chief defects of the Federal statistics—the classification by statutory net income and the large proportion of persons not required to file returns—are remedied by the state statistics, which are classified by total income less capital gains and losses, and which cover a larger proportion of the population. Delaware, in fact, requires returns from every adult, irrespective of the size of his income.

Wisconsin and Delaware are perhaps more prosperous than the average state. Since, however, one is mainly agricultural and the other mainly industrial, they are not likely to deviate in the same direction from the national average in the matter of the composition of income within a given income class. For example, it is not likely that both would show a smaller proportion of the income of the $1,500–$2,000 class derived from labor, or a larger proportion from property, than the average of all the other states, although they might both show that a larger or smaller proportion of all persons fell within the $1,500–$2,000 class than would be the case in other states. For families with incomes under $2,500 and single individuals with incomes under $1,000, conclusions derived from these state statistics are more likely to be correct than those derived from the Federal statistics, since the coverage is better and there is less bias on account of filing requirements and other legal technicalities.

From these statistics it is possible to calculate the proportion of persons in each income class who received various kinds of income, and the proportion of their income received from each of the sources. Applying these proportions to the estimated numbers of persons and of income in each class we can build up aggregates showing the distribution of each

kind of income, hence the ownership of various kinds of capital. These figures can be worked out separately for families and single individuals; in some cases separately for single men and women.

The family figures can be supplemented by the results of the Consumer Purchases Study, which were published in a large number of bulletins by the Bureau of Labor Statistics and the Bureau of Household Economics, and summarized to some extent in the National Resources Committee volumes called *Consumer Expenditures* and *Family Expenditures*. Some of these data refer only to non-relief families and may therefore overstate the importance of some forms of property income in the lower income classes. But if corrected, as they can be in many cases, for relief families, they may tend to understate property incomes, because they were derived to a large extent from families that either were poorer than the average or else understated their income from property.[1]

The CPS data do not state separately profits from individually owned enterprises such as farms, although they make a general distinction between farmers (some of whom earned some wages) and wage earners (some of whom made some profits). Delaware is not so much of a farming state as most, and in addition farmers were believed to be most likely to neglect to file returns. Also in all states business and professional men (including farmers) are generally supposed to be most likely to understate their incomes, as there is less check on net profits than on other forms of income. Consequently the ratios derived from tax returns probably underestimate the importance of profits in all income classes except those large enough to cause the Treasury to make a special audit. No correction on this account has been possible. Wages, partnership income, rent and interest were reported at the source if they exceeded the amount of the individual or family exemption, and are therefore fully represented in the Federal individual tax returns above say $3,500 for families and $2,000 for single individuals. In Delaware they were reported at the source if in excess of $1,000, irrespective of the recipient's family status.

Dividends amounting to $100 or more in a year were reported at the source under the Federal law and were therefore fully reported by persons filing returns, but not necessarily reported at all by families under

[1] For evidence on this point see my articles in the *Review of Economic Statistics*, November 1940 (Vol. XXII, pp. 165–182) and February 1942 (Vol. XXIV, pp. 9–21). Further evidence of understatement (or omission) is found in the fact that the number reporting savings bank accounts far exceeded the number reporting the receipt of either interest or dividends, in every income class. Also, in the income classes above $2,500, where the coverage of the tax returns is complete, the proportion of families reporting dividends and the proportion reporting interest on tax returns, each exceeded the combined proportion of families reporting either dividends or interest on the CPS returns. Likewise the proportion of total income derived from dividends or interest reported on the tax returns each exceeded the proportion of dividends and interest together reported in the CPS, for every income class.

$2,500 or single individuals under $1,000. Consequently a large amount of dividends is omitted, since it was received by persons in the lower income brackets who filed no Federal returns.

In general the tax ratios may slightly exaggerate the importance of property income and minimize the importance of individual profits in all classes between $2,500 and $20,000. Below $1,000 one might suspect that they exaggerated the importance of property incomes, but later on in this article evidence is given to show that that was not the case with respect to the Delaware ratios which have been used for family cash income in classes below $2,500. To these ratios the NRC estimates of imputed rental from owned homes and home-produced food and fuel have been added, since these items were not included in the Delaware returns.

Of course the ratios from the various tax tabulations do not agree exactly; in some cases there are wide variations. But these variations can usually be explained by differences in the legal definition of the kind of income under consideration, or of the income taken as the basis for tabulation. Or else they are caused by differences in the extent of coverage, or by the inclusion of separately reporting husbands and wives in their individual income classes and not in the class where the family belongs. Fortunately it has been possible in every calculation to eliminate the separately reporting husbands and wives, and in most cases to put the combined family in its proper class. If, after this adjustment was made there were two ratios that seemed equally reasonable in the light of the definitions and coverage, that one was taken which showed the lower proportion of property income below $2,500, and the higher proportion above; and the opposite choice was made for labor income.

Besides the tables showing the ratios of all persons receiving income from the different sources, and the total amount received, there are official tables of "Patterns" of income, showing the number of persons wholly relying on income from each source. These "Patterns" are in many respects unsatisfactory; they all lump together families, single individuals and separately reporting husbands and wives. The United States "Patterns" includes estates and trusts. The Wisconsin "Patterns" has excessively wide limits to its income classes. Both the United States and the Wisconsin "Patterns" are tabulated according to a different definition of income from that used in the other tables. The United States "Patterns" tabulates the sources in such a way as to conceal the nature of the fourth source of income and to conceal the number of those having five or more sources. Wisconsin conceals the nature of the third source of income and the number of those having four or more sources. None of the "Patterns" include any imputed income, except an infinitesimal amount of imputed profits in Wisconsin.

Consequently the number of incomes derived wholly from labor, or profits, or property, or profits and property together, is exaggerated by these "Patterns." A considerable amount of correction is possible, by using other tax tables and those in the Consumer Purchases Study, but the figures used in this study probably still overstate the number of single-source incomes, especially single-source property incomes.

The estimate of income distribution to which these ratios were applied was based on the one I published in this JOURNAL last December, but adjusted to the calendar year 1936. If the National Resources Committee's distribution were used as a base it would increase both the number of property incomes and the number of labor incomes in the classes below $1,000 and above $7,500, and would give a larger total amount of property income and a larger number of persons receiving only property income. As I stated in that article, I believe that both the NRC figure and my own, for the under $500 class, were too large.

The definition of income used for classification is the same as that used by the National Resources Committee in its studies, with the inclusion of imputed income for families as estimated by the NRC, but no allowance for imputed income of single persons. There is, therefore, a slight exaggeration of the importance of labor income in the lower classes. Capital gains are not included in income in this analysis.

We may now proceed to a consideration of the detailed tables, and the more important conclusions to be drawn from them.

LABOR (WAGES AND SALARIES)

In the first place the commonest source of income is labor. That is not a startling discovery, but the distribution of labor income by classes is perhaps a little unexpected. The greatest proportion of labor income is found, not in the lowest class (under $500) but in the class $1,000–$1,500, where 90 per cent of the consumer units report it. The proportion declines with increasing income until 71.7 per cent is reached in the $10,000–$15,000 class, but above $15,000 over 75 per cent report labor incomes.

Taking the proportion of labor income to all income: that also is highest in the $1,000–$1,500 class and is higher in every class up to $3,500 than it is in the lowest class (under $500), where it is only 65.7 per cent. Even above $20,000 the proportion of labor income was 22 per cent. Except in the $1,000–$2,000 class, families were more likely than single persons to have labor incomes, and in every class single men were more likely than single women to have labor incomes.

Considering now the proportion of consumer units wholly dependent on labor income, here again the lowest class does not show the highest

TABLE I

PER CENT RECEIVING ANY INCOME FROM SPECIFIED SOURCES, UNITED STATES, 1936
(Consumer units)

Income class	Labor	Business (cash)	Partnership	Imputed profits	Any profits (min.)	Any property or profits (min.)
0– 500	78.2	7.6	.4	15.6	17.8	34.3
500– 1,000	88.2	8.0	.4	18.0	21.4	24.4
1,000– 1,500	90.0	8.4	1.7	16.0	22.3	24.6
1,500– 2,000	89.0	13.7	2.1	14.1	23.8	36.3
2,000– 2,500	86.7	19.0	2.8	11.5	23.2	54.2
2,500– 3,000	83.6	19.2	4.2	12.1	25.7	56.3
3,000– 3,500	81.4	21.4	5.7	11.7	28.0	64.3
3,500– 4,000	80.3	23.7	6.8	11.9	28.9	75.5
4,000– 4,500	79.1	26.7	7.8	10.8	33.6	78.1
4,500– 5,000	77.9	28.6	7.6	10.6	35.9	81.2
5,000–10,000	75.0	34.5	10.5	10.7	46.2	89.2
10,000–15,000	71.7	40.7	13.9	7.0	51.6	96.5
15,000–20,000	75.7	44.6	14.7	7.0	56.5	96.9
20,000 & over	75.3	50.8	18.3	6.9	58.5	99.8

Income class	Fiduciary & other	Interest	Rent (cash)	Imputed rent	Dividends	Any property (min.)
0– 500	.4	17.6	6.7	17.6	7.4	26.8
500– 1,000	.4	10.0	5.0	19.5	5.3	19.5
1,000– 1,500	.5	13.1	5.0	24.6	7.1	24.6
1,500– 2,000	.7	19.4	7.0	31.0	10.6	31.0
2,000– 2,500	1.1	22.2	10.8	38.5	18.7	43.3
2,500– 3,000	1.8	27.2	13.9	47.6	20.4	52.0
3,000– 3,500	2.1	33.3	17.7	47.7	25.0	52.9
3,500– 4,000	2.6	34.2	21.9	46.8	31.9	61.4
4,000– 4,500	3.6	35.2	23.9	55.0	38.0	66.1
4,500– 5,000	4.2	36.5	24.5	55.5	42.1	63.6
5,000–10,000	7.5	44.1	27.4	54.4	47.3	74.9
10,000–15,000	11.1	56.3	28.5	48.9	71.0	83.6
15,000–20,000	13.5	62.4	28.1	48.0	78.9	91.4
20,000 & over	29.3	74.1	32.7	47.2	87.3	95.2

proportion, but is surpassed by the classes between $500 and $1,500. The highest proportion is in the class $500–$1,000, if cash income alone is considered, and in the class $1,000–$1,500 if allowance is made for imputed income. In most classes families appear to be more likely than single individuals to depend wholly on labor income. Perhaps that is because young unmarried persons can save more easily than married ones, and aged persons who have acquired property during their prime are likely to be widowed. In every class single men are much more likely than single women to be dependent on labor only.

If the ratios presented in Table I are applied to the distribution of income presented in my article in the December JOURNAL, the number of

consumer units receiving some labor income comes out as 35,433,000. This represents a considerably larger number of workers since working wives and children are not counted. The total of gainfully employed reported by the 1930 Census was 36,470,000 after deducting farm owners and operators and the unpaid members of their families. Plainly the number of hired workers has not been underestimated in this study.

Similarly the total of wages and salaries obtained by using the ratios of income shown in Table III is $45,387,000,000, while the amount calculated for 1936 by the Department of Commerce is $42,125,000,000.

TABLE II

PER CENT RECEIVING ANY INCOME FROM SPECIFIED SOURCES, UNITED STATES, 1936

Income class	Families*			Single persons†		
	Labor	Profits (min.)	Property (min.)	Labor	Profits (min.)	Property (min.)
0– 500	83.5	30.9	29.5	75.0	3.4	26.8
500– 1,000	89.0	32.4	29.1	87.0	4.9	13.4
1,000– 1,500	90.0	29.0	33.4	92.6	3.8	18.4
1,500– 2,000	88.3	27.2	36.9	90.0	5.7	29.1
2,000– 2,500	86.7	25.4	44.6	84.7	9.4	43.3
2,500– 3,000	82.0	27.1	52.0	80.1	12.8	53.3
3,000– 3,500	78.8	29.3	52.9	74.7	16.2	64.5
3,500– 4,000	76.3	30.5	52.9	63.7	17.6	64.5
4,000– 4,500	75.0	35.6	62.8	61.1	18.8	75.3
4,500– 5,000	77.4	38.2	62.8	57.8	19.1	75.3
5,000–10,000	73.8	50.2	69.1	54.8	20.2	82.3
10,000–15,000	73.6	51.6	82.0	55.1	18.5	92.6
15,000–20,000	77.4	56.5	90.5	53.3	18.4	86.9
20,000 & over	81.3	58.5	96.0	52.7	13.2	94.2

Income class	Single men†			Single women†		
	Labor	Profits (min.)	Property (min.)	Labor	Profits (min.)	Property (min.)
0– 500	87.8	4.2	13.8	64.7	2.7	34.4
500– 1,000	91.6	5.6	8.6	82.2	4.1	18.5
1,000– 1,500	93.2	4.5	12.6	87.6	3.1	25.5
1,500– 2,000	92.8	6.0	20.7	80.6	5.0	42.5
2,000– 2,500	87.3	9.8	29.4	74.3	8.4	60.3
2,500– 3,000	84.0	13.9	41.7	65.1	10.8	67.4
3,000– 3,500	78.1	18.3	56.9	50.0	12.5	73.0
3,500– 4,000	78.1	20.0	56.9	50.0	14.0	73.0
4,000– 4,500	82.0	22.5	66.0	31.9	13.5	85.1
4,500– 5,000	82.0	23.1	66.0	31.9	13.5	85.1
5,000–10,000	69.5	25.5	80.0	20.9	13.7	84.9
10,000–15,000	55.1	24.4	88.9	5.3	12.3	96.3
15,000–20,000	80.0	25.6	80.0	27.3	11.5	92.1
20,000 & over	53.8	20.5	92.3	17.0	8.0	95.7

* Including imputed rent and profits.
† Excluding imputed income.

TABLE III

PER CENT OF INCOME RECEIVED FROM SPECIFIED SOURCES, UNITED STATES, 1936
(Consumer units)

Economic income class	Labor	Business profits (cash)	Partnership profits	Profits (imputed)	All profits	All profits or property
0– 500	65.7	4.3	.3	10.4	15.0	34.3
500– 1,000	75.6	5.7	.3	8.2	14.2	24.4
1,000– 1,500	79.2	6.4	.4	5.4	12.2	20.8
1,500– 2,000	78.0	8.4	.8	3.9	13.1	22.0
2,000– 2,500	75.0	10.0	1.5	2.7	14.2	25.0
2,500– 3,000	72.0	10.6	2.2	2.6	15.4	28.0
3,000– 3,500	68.7	12.3	2.8	2.1	17.2	31.3
3,500– 4,000	64.4	13.9	3.5	2.0	19.4	35.6
4,000– 4,500	60.4	15.1	4.3	1.4	20.8	39.6
4,500– 5,000	58.2	15.1	4.7	1.3	21.1	41.8
5,000–10,000	47.5	14.9	6.2	.9	22.0	52.5
10,000–15,000	39.4	12.7	7.8	.4	20.9	60.6
15,000–20,000	34.3	11.1	8.4	.3	19.8	65.7
20,000 & over	22.0	5.6	9.5	.1	15.2	78.0

Economic income class	Fiduciary & other property	Interest	Rent (cash)	Rent (imputed)	Dividends	All property
0– 500	.4	6.1	4.6	5.6	2.6	19.3
500– 1,000	.4	2.2	2.3	4.0	1.3	10.2
1,000– 1,500	.5	1.4	1.4	4.2	1.1	8.6
1,500– 2,000	.7	1.3	1.5	4.2	1.2	8.9
2,000– 2,500	1.0	1.8	2.0	4.2	1.8	10.8
2,500– 3,000	1.8	1.8	2.2	4.2	2.6	12.6
3,000– 3,500	1.8	2.0	2.7	4.2	3.4	14.1
3,500– 4,000	2.6	2.5	2.7	4.1	4.3	16.0
4,000– 4,500	3.6	2.9	2.6	4.0	5.7	18.8
4,500– 5,000	4.2	3.2	2.6	4.0	6.7	20.7
5,000–10,000	5.4	9.2	2.5	3.5	9.8	30.5
10,000–15,000	8.0	11.3	1.5	3.1	15.8	39.7
15,000–20,000	9.1	12.3	1.5	2.9	20.1	45.9
20,000 & over	11.4	11.7	1.8	2.5	35.5	62.8

The total receiving no cash income except wages or salaries was about 25,000,000 according to these calculations; if an allowance were made for rent of owned homes and imputed profits received by families the total might be as low as 11,000,000. For comparison the total number of consumer units in the United States in 1936 is estimated to have been 41,234,000. But, as will be shown later, these persons with no income but wages and salaries cannot be regarded as a propertyless proletariat. Many of them had high incomes, and most of the families with low incomes owned property, such as automobiles, radios and other durable consumers' goods, and insurance policies, and bank accounts. Moreover many of the single persons owned their homes al-

though it has been impossible to include their imputed rental in these tables.

PROFITS, OR ENTREPRENEURIAL INCOME

Business and partnership profits include the income of farmers, professional people, and unincorporated merchants and manufacturers. They stand between wages and salaries, which represent payment for labor, and the various kinds of property income which are payments for the use of capital. They are a joint product of labor and capital; the persons who receive them are both laborers and capitalists. Individual

TABLE IV

PER CENT OF INCOME RECEIVED FROM SPECIFIED SOURCES, UNITED STATES, 1936

Income class	Families*			Single persons†		
	Labor	Profits	Property	Labor	Profits	Property
0– 500	61.6	23.1	15.3	75.0	3.4	21.6
500– 1,000	71.3	19.3	9.4	85.3	4.3	10.4
1,000– 1,500	76.8	15.0	8.2	90.4	3.1	6.5
1,500– 2,000	77.2	14.6	8.2	84.5	4.5	11.0
2,000– 2,500	75.0	15.4	9.6	73.8	7.0	19.4
2,500– 3,000	73.0	15.8	11.2	62.3	9.5	28.2
3,000– 3,500	70.0	17.7	12.3	57.3	11.9	35.7
3,500– 4,000	66.0	20.1	13.9	47.1	12.9	41.1
4,000– 4,500	62.4	21.6	16.0	40.5	13.6	45.9
4,500– 5,000	60.6	22.0	17.4	38.0	14.0	48.0
5,000–10,000	50.2	23.2	26.6	32.7	13.3	60.2
10,000–15,000	42.6	22.4	35.0	27.0	11.3	70.6
15,000–20,000	37.8	21.3	40.9	23.7	10.9	74.2
20,000 & over	24.6	16.6	58.8	19.5	8.0	81.7

Income class	Single men†			Single women†		
	Labor	Profits	Property	Labor	Profits	Property
0– 500	86.4	4.2	9.4	64.7	2.7	32.3
500– 1,000	90.0	5.3	4.7	80.1	3.4	16.4
1,000– 1,500	92.5	3.7	3.8	87.6	1.8	10.6
1,500– 2,000	89.2	5.0	5.8	74.7	3.0	22.3
2,000– 2,500	82.0	7.9	10.1	55.4	4.9	39.7
2,500– 3,000	73.1	11.3	15.6	41.7	6.2	52.2
3,000– 3,500	64.7	15.0	20.3	31.0	6.5	62.5
3,500– 4,000	59.6	16.4	24.0	24.7	7.4	67.9
4,000– 4,500	54.3	18.0	27.7	20.6	7.2	72.2
4,500– 5,000	52.4	18.8	28.8	17.3	7.2	75.5
5,000–10,000	40.4	19.8	39.8	9.3	6.1	84.6
10,000–15,000	30.9	18.5	50.6	4.9	4.6	90.5
15,000–20,000	26.8	19.3	53.9	3.7	4.2	92.1
20,000 & over	17.3	14.2	68.5	2.5	3.1	94.4

* Including imputed rent and profits.
† Excluding imputed income.

TABLE V

PER CENT OF CONSUMER UNITS RECEIVING INCOME ONLY FROM SPECIFIED SOURCES, UNITED STATES, 1936

Income class	Labor			Profits		Property		
	min.*	min.†	max.*	min.†	max.†	min.*	min.†	max.†
0- 500	28.5	61.7	65.7	3.7	3.7	0.2	15.8	15.8
500- 1,000	41.4	73.8	75.6	4.9	5.0	—	4.2	5.8
1,000- 1,500	34.9	66.1	69.1	4.5	5.2	—	1.7	3.4
1,500- 2,000	21.9	57.1	63.1	5.9	6.5	—	1.5	2.9
2,000- 2,500	2.9	42.0	45.8	6.7	7.1	—	2.0	3.0
2,500- 3,000	—	42.8	43.7	6.3	8.4	—	2.0	4.1
3,000- 3,500	—	29.9	35.7	5.3	8.5	—	3.0	4.4
3,500- 4,000	—	24.5	24.5	5.3	9.1	—	3.0	6.3
4,000- 4,500	—	17.1	21.9	5.7	9.4	—	4.1	6.9
4,500- 5,000	—	17.1	17.8	5.7	9.2	—	4.1	7.8
5,000-10,000	—	5.4	10.8	3.0	8.1	—	5.6	10.4
10,000-15,000	—	—	3.5	.7	5.5	1.6	8.5	16.0
15,000-20,000	—	—	3.1	—	4.2	1.5	8.4	15.4
20,000 & over	—	—	0.2	.9	2.3	4.6	11.5	16.7

* Allowing for imputed rent and profits. The minimum for profits, on this basis, is zero in all classes.

† Cash income only.

business cash profits and partnership profits, as shown by Table I, were proportionally most numerous in the classes above $3,500. But items of imputed profits (food, fuel and merchandise from stock) were commonest in the classes below $3,500. The ratio of all profits to total income, as reported on tax returns, varied slightly among the income classes, and was apparently highest between $3,000 and $15,000. According to the NRC, however, independent professional and business families (including farm families) were most frequent in the class below $500 and the classes over $4,000. Families were more likely than single individuals to receive profits, perhaps on account of the high proportion of married men among farmers. Single men also surpassed single women in the matter of profits, in every income class.

Only a small proportion of consumer units relied wholly on profits. That proportion was greatest in the classes between $1,500 and $4,000, and consisted mainly of families, although according to the Delaware returns a fairly high proportion of single men in the classes between $2,500 and $15,000 may have had no other kind of cash income.

PROPERTY INCOME

Property income is classified as dividends, interest, rent (including a small amount of royalties), and fiduciary income (i.e., income from estates and trusts). These will be taken up in that order, after which property owners of all sorts will be combined.

Dividends.—The total number of consumer units receiving dividends in 1936 was apparently about 4,731,000. Of course there were many other stockholders who received no dividends and in many families there was more than one stockholder. The total number of voters receiving dividends was over 8,200,000, allowing two to a family and assuming that each individual taxpayer is a voter. Dividend receivers were least numerous, proportionately, in the $500–$1,000 class, where only one person in nineteen received dividends; most numerous in the wealthiest class, where over 87 per cent received them. About half of the dividend receivers had total incomes from all sources amounting to less than $2,000, and they received about 10 per cent of the dividends. Of all kinds of income studied dividends were the least evenly distributed. Half of them were received by persons with incomes over $16,000. In the classes under $15,000 single women were much more likely to receive dividends than were single men, and families were less likely.

Interest.—A surprisingly large number of persons in the lowest income brackets, and of single women in all income brackets, received interest. Half of the interest receivers had total incomes under $1,615, and half of the interest income was received by consumer units with total incomes under $3,280. But if the survey made by the Consumer Purchases Study can be relied on there were apparently also a large number of families who had savings accounts, presumably paying interest, but who did not admit receiving any interest, either to the CPS investigators or on their tax returns. However, without counting these owners of non-interest paying savings bank accounts, or attempting to explain them, there were 4,434,000 families and 2,086,000 single individuals who received interest, or over 11,000,000 voters. Including the savings accounts there were 6,282,000 families, making the number of voters 14,650,000. The total amount received appears to have been $2,463,000,000.

Rents.—Rent is a form of income that is nearly as evenly distributed as wages and salaries, if the imputed rent received by home owners is included. There were in 1930, according to the Census, 14,002,000 home owners; in 1940 there were 15,196,000. Only 11,415,000 home-owning families can be located by the methods used in the present study, using the ratios of home ownership in each income class reported by the Consumer Purchases Survey. Half of them had incomes under $1,570. The tax reports enable us to locate 3,282,000 consumer units receiving cash rent, most of whom were probably also home owners, and all of whom have been treated as home owners in this analysis. But since some of them were not, the total number receiving rent, cash and imputed together, is underestimated. Cash rent was more evenly distributed than

dividends, but not as evenly as interest. Half of the consumer units receiving it had incomes under $1,760, and half of the cash rent received was received by consumer units with incomes under $2,250.

Fiduciary (Estates and Trusts).—Fiduciary income was unimportant in amount and received by only a very small proportion of persons in each income class except the highest. In this study it has been necessary to lump fiduciary income together with some miscellaneous sources, such as annuities and other items partaking of the nature of interest but not reported as such on the tax returns.

All Property Income Combined.—When property incomes are lumped together the number in the 0–$500 class, and the amount, are surprisingly large. Checking the Delaware tax data against the CPS figures for families we find that the tax returns show a larger amount of interest and dividends than the CPS; in fact, they show a larger amount of either interest or dividends than the sum of the two as reported by the CPS. But there is reason to believe that the CPS proportions were below the truth.[2]

In the case of rent, other than imputed rent from owned homes, the CPS ratio of families receiving it was higher than the tax return ratio for classes between $500 and $1,500, although the ratio of rent income to total income was lower.

The CPS also reported an item of income other than earnings or rent or dividends or interest or imputed rent. This consisted mainly of annuities and pensions, which are of course property income, although of a non-transferable kind, since they are not paid for current services. Adding together all the property income reported by the CPS the total is a larger proportion of total income than that derived from the tax returns in each class under $2,000; and like the proportions derived from the tax returns they become larger as the income declines from $1,250 toward zero.

A similar increase is shown by the CPS study in the proportion of families receiving cash rent or interest and dividends or imputed rent, as the total income declines below $750. A little reflection shows why this must be so: a large proportion of the people in the lowest income class are there because they are incapable of earning anything—because of age or physical disability. Such small receipts as they may enjoy from property they have saved or inherited constitute a large proportion of their small incomes. In many cases they are able to spend more than their income by gradually drawing down their capital.

The CPS figures for imputed income from home ownership are included in property income in these tabulations. All of the tax return

[2] See footnote 1.

ratios used are based on total income, excluding capital gains, and where necessary further adjusted to allow for imputed income, which was taken from the NRC reports.

THE PROLETARIAT AND THE PROPERTIED CLASS

The simple picture of a class society constructed by Karl Marx and accepted by unobserving persons in this country is not supported by the income tax statistics. We do not have a large propertyless proletariat, living in poverty and despair, ruled by a small class of capitalists, living in luxury and idleness. Every income class received over one-fifth of its income from sources other than wages; no income class received less than one-fifth of its income from the sale of its labor to employers; and no income class received less than 37 per cent of its income from labor of some sort—either working for others or actively conducting a business enterprise. The number of idle rich is small. Of the 231,000 consumer units with incomes over $10,000 not over 51,000 lived entirely on income from property and at least 16,000 or 17,000 of these were single women (including widows).

Of the 5,915,000 consumer units (if that is the right number) with incomes under $500 on the other hand, 935,000 were entirely dependent on property for their income, except if they were able to raise part of their own food and fuel; and of these some 370,000 were single women. One-half of the persons who received any income from property had total incomes under $1,110. The median income of all consumer units receiving wages or salaries was $1,220, while the median income from all sources was $1,214. Another important fact, not shown by the statistics but known to all of us, is that most of the male capitalists had earned their capital by their own efforts, and many of the propertyless laborers will earn enough capital, before they become too old to work, to maintain them in their old age. As a corollary of this we find a large proportion of persons in the lowest 10 per cent of the population—the "submerged tenth"—who live entirely or mainly on the fruits of their savings. These, and the single women of all income classes, constituted almost the whole of our non-working "leisure class" of capitalists.

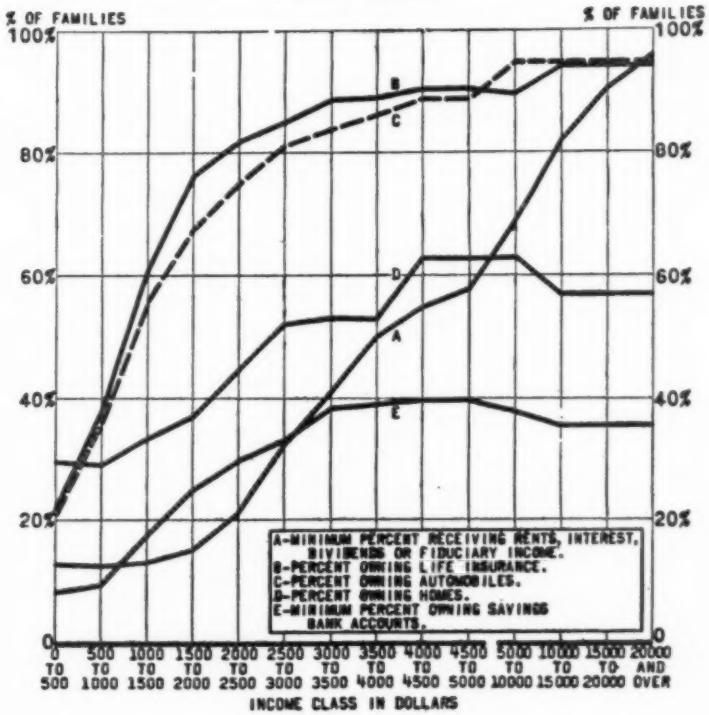
The corporations of the country were not owned by sixty families or 200 large stockholders. It required at least 126,000 of the wealthiest consumer units to muster one-half of the shares. Half of the stockholding units had incomes under $2,030. The unincorporated businesses of the country were even more widely owned; half of the 9,800,000 consumer units receiving business or partnership profits (omitting imputed profits) had incomes under $1,295.

The homes of the country were owned by about 15,000,000 consumer

units, not counting the number indirectly owned through corporations and associations. There were at least 12,700,000 consumer units who owned bonds, mortgages or bank accounts, and half of them had incomes under $1,780.

Some other important forms of capital were not referred to in the tax returns. Some information is available from the Consumer Purchases

PROPORTION OF FAMILIES OWNING PROPERTY
BY INCOME CLASSES
(United States-1936)



Study concerning the ownership of automobiles and insurance policies, and is shown in the attached chart. Clothing and furniture are of course very widely owned but may be regarded as consumer goods rather than capital. But the ownership of a most important type of capital—the tools and machines, shops, stores, inventories and farm equipment necessary to earn business or partnership profits—can only be surmised from the distribution of profits.

On the whole the best measure of the distribution of ownership of all forms of productive capital together appears to be the combined income from property and profits. According to this measure one-half of all

the productive capital of the country was owned by consumer units with incomes under $3,065, and half of the owners had incomes under $1,575.

We must conclude that the vast bulk of our people are both capitalists and laborers during the prime of their lives; they begin their careers without much capital except their bodies and their brains, and they wind up with very little earning power in them, but with enough capital to live on. This fact largely explains the political stability we have had in the past and is the best hope for political stability in the future. Certain other factors in our political stability cannot be measured by the available material, but should be covered by future investigations. They include the age and race factors—to what extent the propertyless persons are young, or Negro, or recently arrived foreigners—and the fluidity factor, by which is meant the shift of individuals from income class to income class and the length of life of large fortunes. My own belief is that, if those factors were studied it would be found that in the matters of equality of opportunity and wide distribution of wealth, the United States compares favorably with any nation now existing, or any important civilized nation of the past.

NOTE: Several tables showing in greater detail the facts mentioned in this article had to be omitted for lack of space. A limited number of copies will be sent by the author to those requesting them.

COMPARISON OF DEMOGRAPHIC RATES ASSUMED BY THE NATIONAL RESOURCES COMMITTEE WITH ACTUAL EXPERIENCE

By Robert J. Myers

In its population projections, the National Resources Committee assumed certain death rates and birth rates for future periods[1] up to 1980. The assumed future rates had as a base point the actual rates for the period 1930–34. Now that five years have elapsed and actual census material is available, it is worth while to make a comparison of the actual rates with those estimated. It is hardly necessary to point out that forecasts of this type are especially dubious for the immediate future because of the effects of the war.

In order to facilitate computations, the type of death rate used by the NRC differed somewhat from the usual age-specific rates of life tables. For a given quinquennial age group the probability of dying within five years was set down for various future quinquennial periods. In order to project births, there was developed a special death rate to give the probability of uniformly-spaced births in a five-year period dying before the end of that time. Three sets of rates, low, medium, and high, were developed for native whites, foreign-born whites, and Negroes with, of course, separate rates for males and females.

Perhaps the best actual mortality data for current purposes is the set of 1940 population life tables[2] prepared by the Statistical Bureau of the Metropolitan Life Insurance Company,[3] from which death rates of the same form as used by the NRC may be developed. In order to make a consistent comparison, it is necessary to make certain adjustments in the NRC figures. First, NRC figures applicable to 1940 may be obtained by averaging the rates for any particular age group for the two quinquennial periods, 1935–39 and 1940–44. Second, a combination of the NRC mortality rates for native whites and foreign-born whites must be made since the Metropolitan life tables which are being used for comparison are only on the basis of total whites. This was accomplished by weighting native white and foreign-born white death rates

[1] The actual figures appear in Tables 1B and 1C of *Population Statistics*, Vol. 1, National Data, National Resources Committee, October 1937. Details of the methodology are contained in P. K. Whelpton, "An Empirical Method of Calculating Future Population," this Journal, September 1936, pp. 457–473.

[2] It will be noted that the actual mortality experience is that for only a single year, 1940. However, the mortality of that year differed only slightly from that of surrounding years; the crude death rate for 1940 was 10.7 per thousand as contrasted with an average of 10.6 for the three years 1939–41.

[3] Presented in abridged form in *Statistical Bulletin*, Metropolitan Life Insurance Company, December 1941, p. 8. The complete table was made available through the courtesy of Dr. A. J. Lotka.

by the respective proportions of these groups relative to the total white population as indicated by the 1940 Census.[4]

In Table I there are set down separately by sex and race the actual 1940 death rates on a comparable basis with the NRC figures. Table IIa and IIb compare these rates with those of the NRC modified as described previously to apply to 1940. At the younger ages the actual rates

TABLE I

ACTUAL 1940 FIVE-YEAR TIME INTERVAL DEATH RATES FOR QUINQUENNIAL AGE GROUPS*

Age group	White persons		Colored persons	
	Males	Females	Males	Females
Birth to 0–4†	48.5	38.1	75.9	59.6
0– 4	10.1	8.3	16.4	13.8
5– 9	5.1	3.7	7.0	5.9
10–14	6.7	4.7	13.1	13.9
15–19	9.8	6.9	25.0	25.0
20–24	12.0	8.9	34.9	31.1
25–29	13.8	10.8	41.8	36.5
30–34	17.5	13.7	49.9	43.8
35–39	24.6	18.1	63.6	57.7
40–44	36.7	25.5	85.3	76.4
45–49	55.4	36.9	117.9	103.9
50–54	82.5	54.1	150.7	136.7
55–59	118.7	80.7	171.2	155.6
60–64	169.5	124.9	209.9	183.0
65–69	243.1	193.8	265.6	230.6
70–74	345.1	289.5	326.5	288.4
75–79	476.5	426.0	412.6	349.0
80–84	623.9	600.0	546.2	420.4
85–89	761.5	774.9	699.7	555.7
90–94	864.9	882.6	832.1	732.6

* Derived from U. S. 1940 Life Tables prepared by the Statistical Bureau of the Metropolitan Life Insurance Company.

† This represents the number of deaths within a five-year time interval among 1,000 infants born uniformly throughout that period.

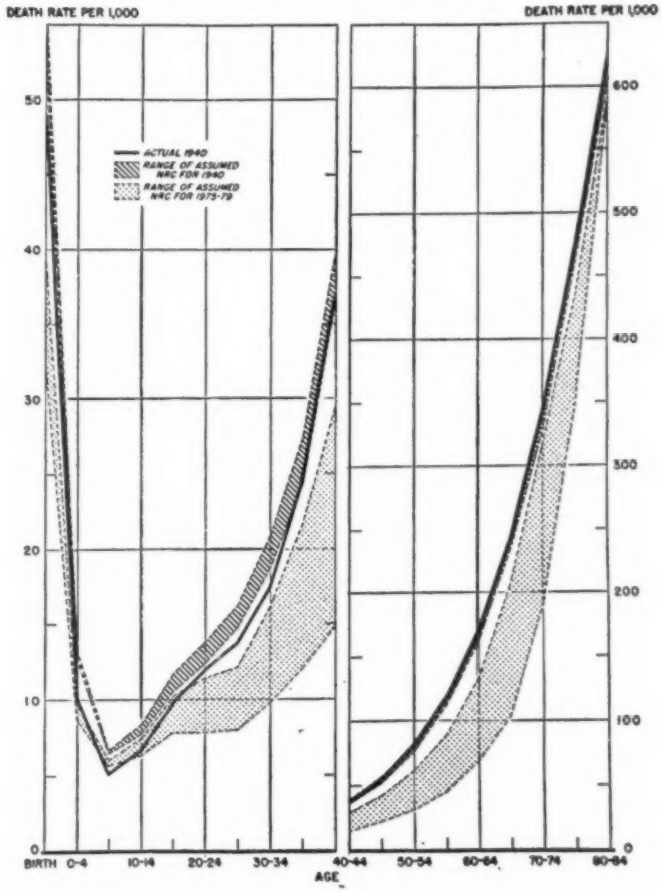
for both white males and white females are appreciably lower than even the "low" estimate; thus, for white males the differential is 20 per cent at the youngest childhood ages and prevails to some extent to age 35, whereas for white females the differential is even larger and persists until the older ages. For white males the actual rates beyond age 35 are about 5 per cent in excess of the "low" estimates, but for white females the corresponding estimates are about 5 per cent too high for ages 35

[4] Based on the release dated April 10, 1942, *Age Composition of the Population*, Series P-10, No. 6, Bureau of the Census.

through 55. Up to age 50 the "high" estimates exceed the actual rates for white males, but beyond there is a slight differential in the opposite direction. The "high" rates for white females are above the actual death rates at all except the extreme ages.

CHART I

COMPARISON OF ACTUAL 1940 FIVE-YEAR DEATH RATES AND THOSE OF NRC—WHITE MALES

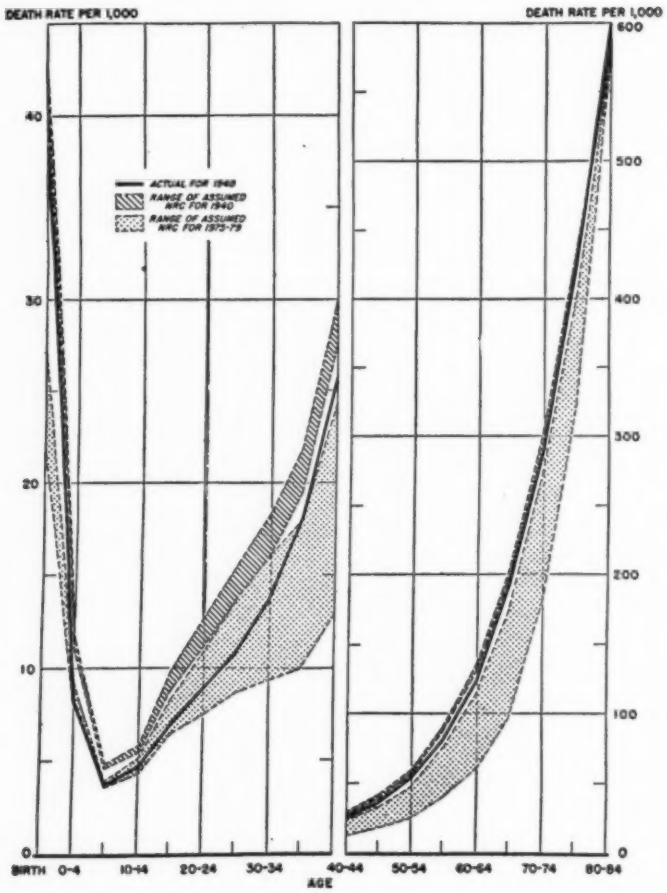


Charts I and II dealing with white males and white females respectively make a graphical comparison of the actual 1940 death rates with those indicated for 1940 by the NRC and also with those for the "ultimate" condition in 1975–79. The difference between the actual rates and the 1940 estimates at the younger ages is quite striking, but be-

yond age 35 for males and 55 for females there is little difference. At the childhood ages for both males and females and at the younger adult ages for females the actual 1940 rates fall well within and, in some instances, below the ultimate range which, it might be pointed out was

CHART II

COMPARISON OF ACTUAL 1940 FIVE-YEAR DEATH RATES AND THOSE OF NRC—WHITE FEMALES



extremely narrow at the younger ages, although wisely wide with advancing age.

Actual 1940 colored death rates have fallen fairly closely within the range developed by the low and high estimate as may be seen from Table IIb. The greater amount of fluctuations present as contrasted

with the fairly smooth trend for white persons in Table IIa arises from the well-known difficulties in obtaining accurate reporting for this population group.

TABLE II

ACTUAL 1940 FIVE-YEAR TIME INTERVAL DEATH RATES AS PERCENTAGE OF ASSUMED NATIONAL RESOURCES COMMITTEE RATES

Age group	a. White persons*				b. Colored persons†			
	Males		Females		Males		Females	
	as per cent of				as per cent of			
	"Low"	"High"	"Low"	"High"	"Low"	"High"	"Low"	"High"
Birth to 0–4‡	94	89	95	90	111	102	107	99
0– 4	80	77	75	72	96	88	89	80
5– 9	80	77	80	77	90	80	95	84
10–14	87	81	87	81	109	91	122	98
15–19	94	84	78	70	107	86	102	81
20–24	97	87	79	70	108	89	108	87
25–29	93	85	79	70	106	91	104	87
30–34	94	84	85	75	98	84	104	87
35–39	101	90	93	83	104	90	108	93
40–44	101	93	93	85	103	93	105	94
45–49	107	99	95	89	114	105	111	102
50–54	109	103	96	91	117	110	105	98
55–59	105	102	92	89	105	101	94	90
60–64	105	101	96	92	103	100	91	87
65–69	104	101	99	96	102	99	95	92
70–74	105	101	99	95	103	99	102	98
75–79	103	101	100	98	94	92	95	94
80–84	103	102	102	102	95	95	97	97
85–89	104	104	107	107	97	97	106	106
90–94	104	104	109	109	104	104	109	109

* Averages of 1935–39 and 1940–44 rates for native and foreign-born whites weighted by respective proportions in 1940 census for given age-sex group (for quinquennial age groups beyond age 75, where census data are not subdivided by age groups, proportion for those 75 and over is used).

† Averages of 1935–39 and 1940–44 rates for Negroes.

‡ See footnote †, Table I.

One other comparison that may be made is to determine the calendar year in which the assumed "medium" death rate is the same as the actual 1940 one (Table III). This will indicate the length of time that has been "gained" in regard to "expected" mortality improvement. For white males up to age 40 there has been a gain of about 5–10 years with the exception of ages 5–15 where we are now already lower than the ultimate figure. For white females there is even more of a gain with such periods being 20–40 years up to age 30 but less thereafter up to age 60. For colored persons there are relatively few gains, except that at the

older ages actual 1940 death rates are in many instances lower than the estimated ones; how much validity may be attached thereto is open to question.

TABLE III

CALENDAR YEAR IN WHICH ASSUMED NATIONAL RESOURCES COMMITTEE "MEDIUM" DEATH RATE EQUALS ACTUAL 1940 RATE*

Age group	White persons		Colored persons	
	Male	Female	Male	Female
Birth to 0–4†	1944	1942	‡	‡
0– 4	1951	Ultimate§	1942	1947
5– 9	Ultimate§	1972	1970	1943
10–14	Ultimate§	1962	‡	‡
15–19	1947	Ultimate§	‡	1942
20–24	1944	1962	‡	‡
25–29	1944	1967	‡	‡
30–34	1944	1955	1942	1941
35–39	1941	1947	1941	‡
40–44	‡	1947	1941	‡
45–49	‡	1945	‡	‡
50–54	‡	1941	‡	‡
55–59	‡	1941	‡	¶
60–64	‡	‡	‡¶	¶

* Rates for quinquennial calendar periods assumed to apply to middle year, with linear interpolation when actual rate falls between assumed ones. Native white rates used for white persons and Negro rates for colored persons.

† See footnote †, Table I.

‡ Actual 1940 rate higher than assumed 1940 "medium" rate derived as described in footnote * of Table IIa and IIb.

§ Actual 1940 rate lower than assumed ultimate (1975–79) "medium" rate.

¶ At most of the ages beyond 55 actual colored death rates are lower than the NRC "medium" estimate for 1940 (see Table II).

The derivation of the actual five-year time interval birth rates was, in general, performed exactly according to the definition which relates births in the given quinquennium (1935–39) to women surviving at the end of the period according to what the age of the mother would have been at that time (regardless of whether or not she actually was alive then). In order to compute such rates, there was first obtained the distribution by single years of age of mother and by race (white and colored) for reported births in each of the years 1935–39.[5]

The births for each of the five years were then collected on a "step" basis so that there was obtained the number of births in the quinquennium to women at each given single age in 1939. For example, the figure

[5] From various issues of *Vital Statistics of the United States*, Bureau of the Census. A slight approximation had to be introduced since single age data are not available for Massachusetts and New Hampshire, but rather only quinquennial values; it was assumed that the distribution within each quinquennial age group was the same for these two states as for the remainder of the country, a procedure which can produce only a negligible error.

for women age 20 in 1939 included births to women age 20 in 1939, age 19 in 1938, etc. Since the distribution of the total female population is known only as of the census data, April 1940, and since the births were related in effect to ages of women as of the middle of 1939, a further adjustment was necessary. Linear interpolation between the births at single years of ages was performed to make this $\frac{3}{4}$ year adjustment. In other words, it was assumed that $\frac{3}{4}$ of the births to women aged x in 1939 applied to women aged $x+1$ as of the census data in 1940, while the remaining $\frac{1}{4}$ applied to women aged x then. Following this subdivision, quinquennial totals were obtained and divided by the 1940 population figures to produce the desired birth rates on a basis unadjusted for under-reporting of births. Such adjustment was then made on the basis used by Whelpton;[1] namely, an under-reporting factor of 96 per cent for white births and 86 per cent for colored births. If reporting accuracy has improved, the finally resulting "actual" birth rates are overstated relative to the NRC assumed rates, but as yet there seems to be no specific evidence on this point.

Table IV shows the actual 1935–39 rates for each quinquennial age group from 10–14 to 55–59, while Chart III deals with the range 15–19 to 45–49 since the NRC did not present figures for the youngest quinquennial age group nor for the two oldest ones, probably because of their small magnitudes. In addition, these actual rates are shown as percentages of the low and high NRC assumptions, with their rates for native and foreign-born whites being blended as was done for death rates and likewise the Negro rates being assumed to apply to colored persons. The total of the actual rates for all ages is indicative of the gross reproduction rate, although not an accurate expression thereof. For white females the actual rates tend to be higher than the low assumptions except for the youngest and oldest age groups but, with only one exception are lower than the high assumptions. Thus at age 25-29 the actual rate is almost 20 per cent in excess of the low assumption and slightly greater than the high assumption. In the aggregate, the actual birth rates are about 10 per cent greater than the low assumptions and slightly lower than the high assumptions. It will be noted that in Chart III there is a considerable margin between the actual 1935–39 rates and the low assumption for 1975–79 (the high assumption for that period being the same as the high assumption for 1935–39).

The actual birth rates for colored females, when contrasted with the NRC assumptions in Table IV, do not show the same smooth relative trends as were present for white females. As with mortality rates, the greater fluctuation undoubtedly arises from inaccuracies and incompleteness of reporting. In the aggregate, the same trends are shown as for white females; namely, an excess of about 10 per cent over the low

assumptions and a slight deficiency as compared with the high assumptions.

In summary, it may be said that our actual mortality rates have improved appreciably more than the National Resources Committee estimated for the younger ages, whereas at the older ages there appears to be some indication of less improvement, especially for white males. In

TABLE IV

ACTUAL 1935–39 FIVE-YEAR TIME INTERVAL BIRTH RATES FOR QUINQUENNIAL AGE GROUPS* COMPARED WITH ASSUMED NATIONAL RESOURCES COMMITTEE RATES†

Age group	White females			Colored females		
	Actual rate	As per cent of "Low"	As per cent of "High"	Actual rate	As per cent of "Low"	As per cent of "High"
10–14	.1	‡	‡	1.6	‡	‡
15–19	55.1	91	85	201.3	129	115
20–24	430.7	106	99	716.3	101	91
25–29	616.7	117	103	675.7	121	106
30–34	496.8	110	97	492.4	118	105
35–39	312.6	104	90	304.8	87	75
40–44	163.9	100	87	209.3	118	101
45–49	42.6	97	84	55.5	97	82
50–54	3.1	‡	‡	8.3	‡	‡
55–59	.1	‡	‡	.6	‡	‡
Total	2121.7	109	97	2665.8	110	97

* Derived from recorded birth statistics for 1935–39 and the 1940 census as described in text.

† Native and foreign-born white rates weighted by respective proportions in 1940 census to give total white rates. Negro rates assumed to apply to total colored persons.

‡ National Resources Committee did not present figures for this age group.

any future projections it would seem advisable to introduce an even more drastic reduction in mortality rates at ages up to 35 (disregarding the immediate and near-future effects of the war, particularly upon males). In this connection it may be noted that the 1941 Hypothetical Life Table,[6] which indicates the plausible long-range future level of mortality and was developed from a corresponding table prepared in 1933, made changes in mortality rates only at ages under 50.

Our actual birth rates have been slightly higher than the medium assumptions of the National Resources Committee but below their high assumptions. As for the main child-producing period, 20–34, the actual birth rates have been about as large as the high assumptions, but at the older ages this tendency was not maintained. A more detailed analysis by child-order would show that this trend arises from the decline in the birth rate for third and higher children.

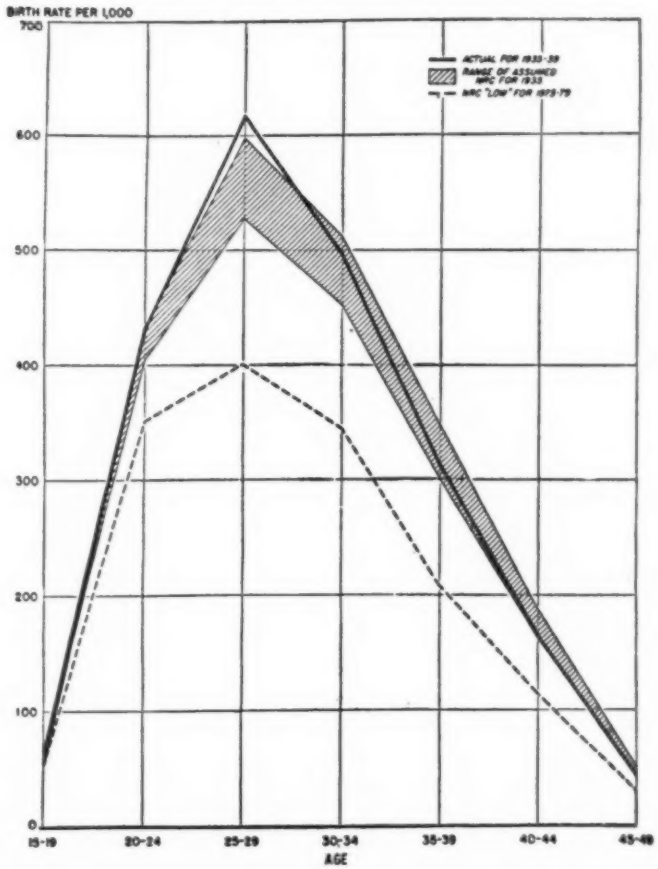
The drastic increase in the birth rate since 1939 would, of course,

[6] *Statistical Bulletin*, Metropolitan Life Insurance Company, November 1941, p. 3.

raise the actual rates considerably above the National Resources Committee assumptions which were based on the birth rates in the trough of the depression, 1930–34. Most authorities are prone to believe that

CHART III

COMPARISON OF ACTUAL 1935–39 FIVE-YEAR BIRTH RATES AND THOSE OF NRC—WHITE FEMALES



the present increase is purely a temporary phenomenon and that after the war age-specific birth rates will continue to decline as is provided in the NRC medium and high assumptions. It is not at all impossible that the present fertility state of the nation is merely a forerunner of what might be expected after the war when economic reconstruction along the lines of the Atlantic Charter will produce social conditions favorable to entirely higher fertility trends.

CHOICE OF THE DEPENDENT VARIABLE IN REGRESSION ANALYSIS

By Frederick V. Waugh
Food Distribution Administration

In the simplest problem of two variables it is always possible to calculate two linear regressions: one with X_1 as the dependent variable, and one with X_2 as the dependent variable. Unless there is perfect correlation between the observed values of X_1 and X_2 these two regression equations will be different. If we plot both on a diagram they will cross each other. How do we know, in any particular case, which regression equation to use?

In a multiple regression problem with n variables we can always calculate n different regression equations treating each variable, in turn, as dependent and the other $(n-1)$ variables as independent. How do we know, in multiple regression, which of these n equations to use?

There appears to be a good deal of confusion about this rather simple matter. A common rule of thumb procedure—particularly among economists—is to select as the dependent variable the one which is taken to represent the "result" or "effect," while the independent variable, or variables, are taken to represent "causes." Those who follow this procedure often argue that only one of the n regressions has any real meaning or any use.

This procedure finds support in some of the best and most recent textbooks. To illustrate this, we shall make several references to the recently revised edition of Dr. Ezekiel's excellent book.[1] Similar references might be found in other textbooks.[2]

Dr. Ezekiel[3] writes,

> It is always necessary to think over the problem carefully, and decide which variable we are going to regard as the independent or causal variable and which one as the dependent, or resultant. Thus, if we were relating variations in tobacco yields to applications of fertilizer, obviously the differences in fertilizer would be the cause and differences in yield the result, so we would sort our records according to the differences in fertilizer.

The last clause might just as well have been written, "so we must use yields as the dependent variable." Dr. Ezekiel states here that we should *always* choose the dependent variable from considerations of cause and

[1] Mordecai Ezekiel, *Methods of Correlation Analysis*, 2nd edition, John Wiley & Sons, New York, 1941.

[2] For example, Albert E. Waugh, *Elements of Statistical Method*, McGraw-Hill Book Company, New York, 1938. See headings, "Nature of Multiple Relationships" and "Dependent and Independent Variables," on pps. 318–319.

[3] Ezekiel, *op. cit.*, pp. 50–51.

effect, but in the following paragraph he indicates that this procedure is not always possible. Thus, he writes:[4]

> It is frequently impossible to say which variable is the cause and which is the effect. All that can be definitely established is that the two vary together. Yet one may wish to regard one variable as the one whose values are given or known. It is then called the independent variable. . . . The second variable will then be regarded as the one whose values are to be related to, or estimated from, the values of the known variable. It is then called the dependent variable since it is treated as depending upon the given values of the independent variable.

These two quotations appear to summarize Dr. Ezekiel's views on this problem. He seems to advise us to follow the cause and effect principle whenever we know which is which. When we don't know which is cause and which is effect, or when both variables are results of a common cause, we are advised to use another criterion in choosing the dependent variable. Not only does Dr. Ezekiel advise us to follow this procedure, he follows it consistently himself in all the problems presented in his book. Whenever one variable is clearly the result, it is chosen as the dependent, and in such cases Dr. Ezekiel seems to feel that only this one regression equation has any meaning or any use. It is largely on these grounds that Dr. Ezekiel recommends in his technical appendix[5] a computational method which gives only one of the n multiple regression equations, when it would be possible with very little additional work to get them all.[6]

What about this procedure? Does it give us the regression equation we want in each particular case?

Suppose we take the example given by Dr. Ezekiel: tobacco yields and fertilizer applications. Let us imagine that we have surveyed 100 farms in North Carolina, and have found the average yield of tobacco and the rate of fertilizer application on each farm. We want a line or equation which will express the relation between the two variables: tobacco yield and fertilizer application. There are two such lines (neglecting the possibility of choosing between different types of curves)—one with yield dependent and one with fertilizer dependent. Dr. Ezekiel says that variations in fertilizer are obviously to be considered as the cause of the variations in yield, and, therefore, that yield must be the dependent variable. Granting that the direction of the cause and effect relation is clear in this case, the writer fails to see any reason for con-

[4] *Op. cit.*, p. 51.

[5] *Op. cit.*, Appendix I, pp. 455–485.

[6] Frederick V. Waugh, "A Simplified Method of Determining Multiple Regression Analysis," this JOURNAL, December 1935, and "The Analysis of Regression in Subsets of Variables," this JOURNAL, December 1936.

cluding that yield must be the dependent variable. We might use this set of data to answer either of the following questions:

a. What were the average yields of tobacco associated with various amounts of fertilizer? The answer to this question is provided by the regression with yield as the dependent variable.

b. What were the average amounts of fertilizer associated with various yields of tobacco? The answer to this question is provided by the regression with fertilizer application as the dependent variable.

We shall not prove these statements here. The proofs can be found in practically any textbook on correlation and regression.[7] The point is that *both* regressions have definite meanings and definite uses, and our choice does not depend at all upon which is cause and which is effect, but simply upon the question we want answered. We may very well want both regressions, because we may very well want to answer both questions.

The same point may be illustrated with any problem which may come to the reader's mind. However clear the direction of the cause and effect relation may be there are always two essentially different questions which may be asked in any problem of two variables. If we want the average, or expected, value of X_1 associated with any particular value of X_2, the answer is given by the regression with X_1 dependent—regardless of the nature of cause and effect. If we want the average, or expected, value of X_2 associated with any particular value of X_1, the answer is given by the regression with X_2 dependent—regardless of the nature of cause and effect.

If we want to estimate the average, or expected, height of fathers corresponding to a given height of sons, we should take height of fathers as the dependent, regardless of the fact that this must be the causal factor. If we want the average amount of potato production associated with parity prices of potatoes, we should take potato production as the dependent, even if we are sure that variations in production *cause* the variations in price.

In a big multiple regression problem we often think in terms of one particular regression. Thus, in a farm management survey we may be concerned with a large number of (causal) factors which affect farm income (the resultant). In such a problem our thinking tends usually to run in the direction of investigating what changes in average, or expected, income would tend to be associated with particular adjustments

[7] For example, in Ezekiel, Technical Note 2, on page 489. Note that Y is selected as the dependent variable without any consideration of whether it represents cause or effect. In any case the regression minimizes the sum of squares of the Y residuals. And in any case a similar mathematical development would show that by choosing X as the dependent variable we minimize the sum of squares of the X residuals.

in the several causal variables. We tend to frame our questions in such terms as, "How much increase in income could be expected on this farm if we added 50 chickens?" or "If these farmers fed each of their cows another pound of feed each day, how much would their average income be increased?" Both of these questions are answered best by the regression with income as the dependent variable.

Does this mean that the other $(n-1)$ regressions have no value? Not at all. Let us think of our problem in different terms. Let us suppose that some of the farmers in our survey must increase their incomes by $200 a year to get a minimum standard of living. They hope to do this by raising more chickens, but without making any other adjustments. We want to know then the average number of chickens associated with the higher income and with the same quantities of the other factors. The answer to this is given by the regression with the number of chickens as the dependent, and with income and all the other variables as independents.

Further examples should be unnecessary. The reader can easily apply similar reasoning to *any problem* of related variables, and will find that in each case there will be as many useful regression equations as variables, and that each of these equations provides the answer to a reasonable question.

Now, to summarize, and to add a few general comments.

1. In any problem of relationships between n variables, each of the n regression equations has a definite meaning, and each can be used to answer a specific type of question.

2. Even when our problem is such that one variable is clearly the "result," and the other variables are clearly "causes," we must choose one of the "causal" variables as the dependent if we wish to estimate the average, or expected, value of that variable corresponding with any set of values for the others.

3. We may very often wish to estimate the average, or expected, value of one of the "causal" variables needed to produce a stated "result." In any such case the "causal" variable becomes the dependent. The "resultant" variable, and any other "causal" variables are treated as independents.

4. There is no sound reason for the common rule of thumb practice of always designating current price as the dependent when relating it to such variables as current supplies and current incomes of consumers; nor for always designating consumption as the dependent when it is related to current retail prices; nor for always designating acreage of a crop as the dependent variable when it is related to prices of that crop; nor for always designating the yield of a crop as the dependent when re-

lating it to fertilizer applications and other variables representing cultural practices. This rule of thumb practice is based on the mistaken idea that the only useful regression equation is the one with the "resultant" variable as the dependent.

5. Nothing in this paper implies, in any way or in any degree, that we can neglect careful logical analysis in the statistical measurement of relationships. The writer subscribes completely to Ezekiel's chapter[8] emphasizing the need for such logical analysis. The point is only that considerations of cause and effect do not indicate which variable should be dependent.

6. We are concerned in this paper only with the classical regression equations. These are to be used whenever our purpose is to estimate values of one of the observed variables. If, however, we assume perfect correlation between the true, underlying values of our variables, we are assuming that departures are due to errors of observation, or to what Frisch[9] calls "disturbances." If we can estimate the relative magnitudes of these errors or disturbances in each of the variables it is possible to calculate one line in the case of two variables, or one n-dimensional plane in the case of n variables expressing the relationship between the true, or basic values. In this case there is only one line or plane, and no one variable is taken as dependent. The so-called "orthogonal regression" assumes equal errors, or disturbances, in each variable. No such line or plane gives any measure of correlation, since correlation is assumed to be perfect among the true, underlying values. Neither does such a line or plane provide an estimate of the average, or expected, values of any of the observed variables.

[8] *Op. cit.*, Chapter 24, pp. 442–453.

[9] Ragnar Frisch, "Statistical Confluence Analysis by Means of Complete Regression Systems," Universitets Økonomiske Institutt, Oslo, Norway, 1934, especially Part II, pp. 47–83.

COMMENTS

BY MORDECAI EZEKIEL
United States Department of Agriculture

Dr. Waugh agrees that where we wish to estimate the most probable value of one variable from a given value of another, we take the latter as the independent variable, and the former as the dependent. He then goes on to say, however, that "both regressions have definite meanings and definite uses, and our choice does not depend at all upon which is cause and which is effect, but simply upon the question which we want answered." My own belief is that in cases where we have definite non-statistical knowledge of which variable is cause and which is the effect,

the reverse regression, which measures the average or most probable value of the "cause" associated with given values of the "result," is generally of little significance and is likely to be used for illogical or incorrect conclusions. Where, by the use of multiple correlation, we can eliminate other associated causes and come still closer to a measurement of the influence of each separate cause upon the result, the use of the reverse partial regression is, I believe, likely to be even more misleading.

The misleading character of the reverse regression is shown by Dr. Waugh's own illustration of farmers who wish to increase their incomes by $200 a year by raising more chickens, while holding the rest of their farming unchanged. Dr. Waugh would apparently determine this by finding the net regression of chicken numbers on farm income, while holding constant the other independent variables. Yet, for each individual farmer involved, what he needs to know is not "What numbers of chickens, on the average, are associated with given incomes?" but rather "How many chickens need I add to my farm to increase my income $200?" What the farmer can do is to *change the number of chickens;* that is the *cause* which he will employ to change his income. The most accurate way to find how much *each chicken* adds to income is to employ the net regression of income upon number of chickens. If this comes out 50 cents, that will mean the 400 chickens will be the number probably required. If the question is restated as how many chickens for a given sum instead of how much money per chicken, that does not change the fact that the *chickens* are what the farmer can actually change.

This differs from Dr. Waugh's point of view. He states "we may wish to estimate the average, or expected, value of that variable (chickens) corresponding with any set of values for the others (income)." While "we" may at times wish to know how many hens the $500 farmer had, and how many the $700 farmer had, the farmer has only an academic interest in those facts. What he wants to know is, how much effect will a *change* in hens have upon his income, for the *hens* are the only thing he can vary directly. The income which will follow as a consequence is still unknown when the hen numbers are decided upon.

I do agree that if we were to take new sample farms from the same universe, and to calculate for those new samples the average number of hens on farms having given incomes, the best advance estimate we could make of that relation would be based on the regression of hens on income in our original sample. But I question whether that information is of any value to the farmer who has to decide what he will do with his poultry flock—or for that matter, is often of very much value to anyone else.

There has been much fallacious reasoning based upon the reverse

regression approach. One common illustration is the type of farm management report which shows the average number of acres of each crop and head of each class of livestock on farms *reporting various levels of income*. It has been the evident conclusion that because the men who had high incomes had more dairy cows, it would therefore pay the men who had lower incomes to increase the number of their cows. Yet actually a classification of the same data by dairy cow numbers might show only insignificant changes in average incomes with differences in numbers of cows. Similarly, in the poultry case, because of some other variable not explicitly measured—such as the ability of the more intelligent farmers to handle all their livestock better—the data might show relatively small increases in chickens associated with very large differences in income, perhaps one more hen for each ten dollars more income. Dr. Waugh would then conclude that a small change in chicken numbers was associated with a corresponding large increase in income. This information would tend to mislead the individual farmers, who after all were in position to change their chicken numbers but not to change their own inherent abilities. In fact, the lower the correlation between income and hens, and the smaller the differences in hens associated with given differences in incomes, the *fewer* the number of hens which this method would indicate as needed to increase farm income by $200!

In view of these considerations, I still believe that the procedures recommended in my book are the most satisfactory for the practical research worker who expects to use his research as a basis upon which to advise on the probable consequences of actions in the real world, for in that world, people need to know what effects (results) are most likely to follow from given things they may select to do (causes).

NOTE ON A COMMON STATISTICAL INEQUALITY

By Irving H. Siegel
U. S. Bureau of Labor Statistics

This paper is devoted to the discussion of a family of inequalities frequently encountered in statistical work. The prototype relation is

$$\sum st \sum uv \gtreqless \sum sv \sum ut \tag{1}$$

or

$$\Delta = \begin{vmatrix} \sum st & \sum sv \\ \sum ut & \sum uv \end{vmatrix} \gtreqless 0, \tag{1a}$$

where all of the sums include the same number of items. In the terminology of Hardy-Littlewood-Pólya, the inequality is "homogeneous in $\sum$" as well as in the variables, and the two sides of (1) are not necessarily "comparable." That is, the sense of the inequality is not affected by the substitution of means for the corresponding sums, but it need not be fixed for all non-negative values of the variables.[1]

In the following, we shall first present some useful identities related to (1a) and then analyze some cases of the inequality which are of interest to the statistician. The validity of the identities is independent of the sign of the variables. In some of the inequalities mentioned later, however, one or more of the variables will be restricted to non-negative values. In any case, only real discrete variables are considered; and all of the sums are finite. Analogous inequalities generally exist for integrals and infinite sums.

The factorability of the matrix of Δ into two rectangular ones is a very useful property. If each of the sums contains n elements, we have

$$|\Delta| = \left\| \begin{matrix} \sum st & \sum sv \\ \sum ut & \sum uv \end{matrix} \right\| = \left\| \begin{matrix} s_1 \cdots s_n \\ u_1 \cdots u_n \end{matrix} \right\| \cdot \left\| \begin{matrix} t_1 \cdots t_n \\ v_1 \cdots v_n \end{matrix} \right\|. \tag{2}$$

For our purposes, the rectangular matrices may be considered permutable: the multiplication is performed row by row, and the value of Δ is not altered by the interchange of diagonal terms. The value of Δ also remains unaltered if the order of the rows is reversed in both rectangular matrices, but its sign is changed if the order is reversed in only one. The matrix of Δ may also be decomposed into products of other rectangular matrices; for the sums in Δ may be expressed in terms of new variables (derived from the original ones) without being altered in value or general form.

[1] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities* (1934), pp. 1–6.

By virtue of (2) or any of its variants, Δ may be transformed into a double sum containing the ${}_nC_2$ products of second-order determinants obtainable from the rectangular matrices. For example, it follows from (2) that

$$\Delta = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \begin{vmatrix} s_i & s_j \\ u_i & u_j \end{vmatrix} \cdot \begin{vmatrix} t_i & t_j \\ v_i & v_j \end{vmatrix}. \tag{3}$$

In many special cases of importance, each determinant product is positive (or negative) and Δ is necessarily greater (or less) than zero.

The sum (3) may be regarded as the Cartesian form of a well known vector identity which is often associated with the name of Lagrange. If the s_i, t_i, u_i, and $v_i (i = 1, \cdots, n)$ are the n rectangular coordinates of the vectors $\boldsymbol{s}$, $\boldsymbol{t}$, $\boldsymbol{u}$, and $\boldsymbol{v}$, respectively, then

$$\Delta = \begin{vmatrix} \boldsymbol{s} \cdot \boldsymbol{t} & \boldsymbol{s} \cdot \boldsymbol{v} \\ \boldsymbol{u} \cdot \boldsymbol{t} & \boldsymbol{u} \cdot \boldsymbol{v} \end{vmatrix} = (\boldsymbol{s} \times \boldsymbol{u}) \cdot (\boldsymbol{t} \times \boldsymbol{v}), \tag{4}$$

where a dot signifies a scalar product and a cross signifies a vector product.[2] As will be illustrated later, it is sometimes enlightening to translate (4) into its trigonometric equivalent.

There is often some advantage to be gained by the introduction of a weighted correlation coefficient which always has either the same sign as Δ or exactly the opposite.[3] In either case, the correlation coefficient and Δ vanish identically. For example, we may write

$$\begin{aligned} \Delta &= \sum st \left(\sum uv - \frac{\sum ut \sum sv}{\sum st} \right) \\ &= \sum st \left[\sum st \left(u/s - \frac{\sum stu/s}{\sum st} \right) \left(v/t - \frac{\sum stv/t}{\sum st} \right) \right] \\ &= r_{st:\,(u/s)\cdot(v/t)} \sigma_{st:\,(u/s)} \sigma_{st:\,(v/t)} \left(\sum st \right)^2, \end{aligned} \tag{5}$$

where $r_{st:\,(u/s)\cdot(v/t)}$ is the coefficient of correlation between the u_i/s_i and v_i/t_i with weights $s_i t_i$ and $\sigma_{st:\,(u/s)}$ and $\sigma_{st:\,(v/t)}$ are similarly weighted standard deviations. We may also express Δ in forms analogous to (5) containing the correlation coefficients $r_{uv:\,(s/u)\cdot(t/v)}$, $-r_{sv:\,(u/s)\cdot(t/v)}$, and

[2] W. F. Osgood, in his *Advanced Calculus* (1935), p. 306, follows W. Blaschke, *Vorlesungen über Differentialgeometrie* (1921), I, pp. 6 and 13, in attributing (4) to Lagrange. But Lagrange's contribution, included in a paper presented in 1773, is simply an algebraic identity corresponding to the case of (3) in which $n = 3$, $s_i = t_i$, and $u_i = v_i$. Some writers (e.g., Courant and Goursat-Hedrick) ascribe only this case to Lagrange. See T. Muir, *Theory of Determinants in the Historical Order of Development* (2nd ed., 1906), I, pp. 33–37.

[3] For applications of weighted correlation coefficients, see the following articles by the writer in this JOURNAL: "The Difference between the Paasche and Laspeyres Index-Number Formulas," September 1941, pp. 343–350; "Further Notes on the Difference between Index-Number Formulas," December 1941, pp. 519–524; and "Index-Number Differences: Geometric Means," June 1942, pp. 271–274. See also the papers of L. von Bortkiewicz mentioned in the first of these articles.

$-r_{ut:(s/u)\cdot(v/t)}$. In each instance, Δ is numerically equal to a weighted product moment times the square of the sum of the weights. A useful relation is thus established between product moments and double sums.

To emphasize the symmetry between the correlation and double-sum forms of Δ, we may explicitly introduce into (3) the variables and weights entering into the correlation coefficient in (5):

$$\Delta = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} s_i t_i s_j t_j \left(\frac{u_j}{s_j} - \frac{u_i}{s_i}\right)\left(\frac{v_j}{t_j} - \frac{v_i}{t_i}\right). \tag{3a}$$

Similarly, we may rewrite (3) to correspond to any of the other weighted correlation coefficients appearing in alternative forms of (5). Incidentally, variants of (3) such as (3a) could, of course, be obtained directly from variants of (2) if the sums in $|\Delta|$ were first rewritten appropriately.

We now turn to an examination of some of the inequalities represented by Δ. First, we shall consider instances in which the sense of the inequality is not fixed.

An important example is the index-number inequality, which has already been discussed at some length in this JOURNAL.[4] If the u_i and s_i represent, say, alternative sets of quantity weights and the v_i and t_i represent prices in two periods, the latter being the base, then Δ gives the sign of the difference between two aggregative price indexes. The sign of the difference between the same two indexes converted to weighted arithmetic means of the relatives v_i/t_i is given by $\Delta = \sum ST \sum UV - \sum SV \sum UT$, where $S_i = s_i t_i$, $T_i = 1$, $U_i = s_i t_i (u_i/s_i)$, and $V_i = v_i/t_i$; the two sets of weights are U_i and S_i. This result is widely applicable, for any harmonic mean of relatives is also equivalent to a weighted arithmetic mean. In the special case where the u_i/s_i (i.e., the ratios of the alternative sets of weights for either the aggregative or mean forms) represent quantity relatives corresponding to the price relatives being averaged, we have the difference between the Paasche and Laspeyres indexes. From (3a), it is evident that the Paasche index is necessarily greater than the Laspeyres index if the coefficient of rank correlation between the two sets of relatives is $+1$, and necessarily smaller if the value of the coefficient is -1. Tchebycheff obtained an equivalent result for analogous integrals.[5]

The tetrad difference, $r_{a\cdot b}r_{c\cdot d} - r_{a\cdot d}r_{c\cdot b}$, is a fundamental notion of psychometrics and also occurs in a simple case of vector correlation.[6] It

[4] See papers by the writer mentioned in footnote 3.

[5] *Inequalities*, pp. 43–44.

[6] For the special case of vector correlation having the tetrad difference as numerator, see H. Hotelling, "Relations between Two Sets of Variates," *Biometrika*, December 1936, p. 324, or C. C. Craig,

may be represented by Δ if the variables s_i, t_i, u_i, and v_i are replaced by $a_i' = (a_i - \sum a/n)/\sqrt{n}\sigma_a = (a_i - \sum a/n)/[\sum(a - \sum a/n)^2]^{1/2}$ and analogously defined b_i', c_i', and d_i'. The tetrad vanishes when a weighted correlation coefficient such as $r_{a'b':(c'/a')\cdot(d'/b')}$ or $r_{c'd'(a'/c')\cdot(b'/d')}$ equals zero; or, in geometric terms, when the planes $\boldsymbol{a}'$, $\boldsymbol{c}'$ and $\boldsymbol{b}'$, $\boldsymbol{d}'$ (the vectors have the coordinates a_i', c_i', b_i', and d_i', respectively) are at right angles.

The sign of the partial correlation coefficient $r_{ab\cdot c}$ is given by the sign of its numerator, $r_{a\cdot b} - r_{a\cdot c}r_{b\cdot c} = r_{c\cdot c}r_{a\cdot b} - r_{a\cdot c}r_{b\cdot c}$, which is in the form of Δ for the variables a_i', b_i', and $c_i' = d_i'$ defined as in the preceding paragraph (whence $\sum c'^2 = 1$). Thus, the sign of the partial coefficient corresponds to that of $r_{a'b':(a'/a')\cdot(c'/b')}$ or $r_{c'^2:(a'/c')\cdot(b'/c')}$.

The signs of the constants of the regression equation $Y = f + gX$ may also be considered to depend on expressions like Δ which form the numerators of the determinant solutions to the normal equations.[7] For example, if the observation equations are weighted by w_i, then the following substitutions in Δ give the numerator of the regression coefficient g: $s_i = w_i$, $t_i = 1$, $u_i = w_iX_i$, and $v_i = Y_i$. If the observation equations are unweighted, then $s_i = t_i = 1$, $u_i = X_i$, and $v_i = Y_i$. The corresponding weighted and unweighted correlation coefficients, $r_{w:X\cdot Y}$ and $r_{X\cdot Y}$, have the same signs as these Δ's.

We now turn to instances in which the two sides of (1) are "comparable"—in which the sense of the inequality is fixed for at least non-negative values of the variables. In each case, we shall make the left member of (1) equal to or greater than the right; that is, Δ will be positive except in cases of equality.

A familiar definitely directed inequality is the difference between the (weighted or unweighted) arithmetic and harmonic means of a set of positive quantities. Thus, if the z_i are positive, $\sum z \sum z^{-1} \geqq n^2$ is clearly valid, for the determinant products derivable from

$$\begin{Vmatrix} \sum z & n \\ n & \sum z^{-1} \end{Vmatrix} = \begin{Vmatrix} z_1 \cdots z_n \\ 1 \cdots 1 \end{Vmatrix} \cdot \begin{Vmatrix} 1 & \cdots & 1 \\ z_1^{-1} & \cdots & z_n^{-1} \end{Vmatrix}$$

are all non-negative.

Another common relation which may readily be verified by the methods described in this paper is the following moment inequality for (weighted) positive quantities:

$$\frac{\sum pz^m}{\sum p} \geqq \frac{\sum pz^k}{\sum p} \frac{\sum pz^{m-k}}{\sum p}, \qquad m > k > 0, \quad p_i > 0.$$

---

[6] "Recent Advances in Mathematical Statistics, II," *Annals of Mathematical Statistics*, March 1942, p. 76. For vector treatment of tetrads, see E. B. Wilson, "On Hierarchical Correlation Systems," *Proceedings of the National Academy of Sciences*, 1928, pp. 283-291.

[7] The determinant in the denominator also has the form of Δ but may be shown by (3) to be necessarily positive; in fact, this determinant is an example of Cauchy's inequality, which is discussed several paragraphs later.

Decomposition of the matrix

$$\begin{Vmatrix} \sum p & \sum pz^k \\ \sum pz^{m-k} & \sum pz^m \end{Vmatrix}$$

shows that the typical determinant product is non-negative. This inequality may be extended by repeated application to the factors on the right-hand side.

The validity of the following moment inequalities, in which the weights p_i are positive, may also be established readily by examination of the typical determinant products:

$$\sum p \sum pz^{2m} \geqq (\sum pz^m)^2$$
$$\sum pz^{2k} \sum pz^{2m} \geqq (\sum pz^{m+k})^2$$
$$\sum pz^{m+1} \sum pz^{m-1} \geqq (\sum pz^m)^2.$$

The first and second of these inequalities hold for negative as well as positive values of z_i and of the exponents; the third holds for mixed values of z_i if m is an odd integer.

The most famous special case of (1) is Cauchy's inequality, which occurs not only in a variety of statistical problems but also in the derivation of more complex inequalities, like Hölder's and Liapounoff's. The matrix of Δ for Cauchy's inequality may be written as the product of two identical rectangular matrices; the typical determinant product is a perfect square. All but one of the definitely directed inequalities which have already been considered may be written in Cauchy's form. For example, the expression giving the sign of the difference between an arithmetic and a harmonic mean is equivalent to the following:

$$\sum z \sum z^{-1} - n^2 = \sum (z^{1/2})^2 \sum [(z^{-1})^{1/2}]^2 - [\sum z^{1/2}(z^{-1})^{1/2}]^2 \geqq 0.$$

Cauchy's inequality is equivalent to the statement that the cosine of the angle between two vectors is never greater than unity. Since any correlation coefficient may always be written in cosine form, its value obviously cannot exceed ± 1, as is well known. By virtue of (3) or (3a), various correlation coefficients may be reduced to cosine form in more than one way; that is, the difference between the denominator and numerator may be given by more than one Cauchy inequality. For example, at least two such expressions exist for the weighted zero-order coefficient:

$$r_{w;\,a\cdot b} = \frac{\sum w(a - \sum wa/\sum w)(b - \sum wb/\sum w)}{[\sum w(a - \sum wa/\sum w)^2 \sum w\,(b - \sum wb/\sum w)^2]^{1/2}}$$
$$= \frac{\sum_i \sum_j w_i w_j (a_i - a_j)(b_i - b_j)}{[\sum_i \sum_j w_i w_j (a_i - a_j)^2 \sum_i \sum_j w_i w_j (b_i - b_j)^2]^{1/2}}.$$

In the unweighted case, we have

$$r_{a\cdot b} = \frac{\sum (a - \sum a/n)(b - \sum b/n)}{[\sum (a - \sum a/n)^2 \sum (b - \sum b/n)^2]^{1/2}} = \frac{\sum_i \sum_j (a_i - a_j)(b_i - b_j)}{[\sum_i \sum_j (a_i - a_j)^2 \sum_i \sum_j (b_i - b_j)^2]^{1/2}}.$$

The first-order coefficient $r_{ab\cdot c}$ may be expressed in an alternative cosine form involving the a_i', b_i', and c_i' defined in an earlier paragraph:

$$r_{ab\cdot c} = \frac{r_{c\cdot c}r_{a\cdot b} - r_{a\cdot c}r_{b\cdot c}}{(r_{a\cdot a}r_{c\cdot c} - r_{a\cdot c}r_{a\cdot c})^{1/2}(r_{b\cdot b}r_{c\cdot c} - r_{b\cdot c}r_{b\cdot c})^{1/2}} = \frac{\sum_i \sum_j (c_i'a_j' - c_j'a_i')(c_i'b_j' - c_j'b_i')}{[\sum_i \sum_j (c_i'a_j' - c_j'a_i')^2 \sum_i \sum_j (c_i'b_j' - c_j'b_i')^2]^{1/2}}$$

Partial coefficients of higher order may be written in analogous fashion in terms of multiple sums of products of higher-order determinant pairs.

In conclusion, it may be well to recall at this point the analogy between the first-order correlation coefficient and the fundamental cosine law of spherical trigonometry.[8] Since the latter may be represented by a Lagrange identity,[9] so may the former. Thus, if $\mathbf{a}'$, $\mathbf{b}'$, and $\mathbf{c}'$ are vectors of unit length having the rectangular coordinates a_i', b_i', and c_i', respectively ($i = 1, \cdots, n$; $\sum a'^2 = \sum b'^2 = \sum c'^2 = 1$), then

$$(\mathbf{c}' \times \mathbf{a}') \cdot (\mathbf{c}' \times \mathbf{b}') = \mathbf{c}' \cdot \mathbf{c}' \mathbf{a}' \cdot \mathbf{b}' - \mathbf{a}' \cdot \mathbf{c}' \mathbf{b}' \cdot \mathbf{c}'.$$

The partial coefficient corresponds to the cosine of the dihedral angle between the planes $\mathbf{c}'$, $\mathbf{a}'$ and $\mathbf{c}'$, $\mathbf{b}'$. The scalar products $\mathbf{a}' \cdot \mathbf{b}'$, $\mathbf{a}' \cdot \mathbf{c}'$, and $\mathbf{b}' \cdot \mathbf{c}'$, which represent the cosines of the angles subtended by the sides of the spherical triangle, also correspond to the simple correlation coefficients $r_{a\cdot b}$, $r_{a\cdot c}$, and $r_{b\cdot c}$. The vector identity may readily be translated into the double-sum coordinate form of the partial coefficient shown in the preceding paragraph.

[8] See D. Jackson, "Trigonometry of Correlation," *American Mathematical Monthly*, June 1924, pp. 275–280, or R. W. Burgess, *Introduction to Mathematics of Statistics* (1927), pp. 275–276. For a comprehensive treatment of correlation in general, see J. MacMahon, "Hyperspherical Goniometry; and Its Application to Correlation Theory for N Variables," *Biometrika*, December 1923, pp. 173–208.

[9] See, for example, C. E. Weatherburn, *Elementary Vector Analysis* (1939), p. 72.

CONTENTS OF TIPPETT'S "RANDOM SAMPLING NUMBERS"

By Robert Gage, M.S.
Section on Biometry and Medical Statistics
Mayo Clinic, Rochester, Minnesota

As part of a plan for setting up a series of punch cards to be used for experiments with random samples, there were preliminarily punched onto Hollerith cards all the numbers in Tippett's tables of random sampling numbers.[1] In view of the wide use of this table for random experiments of various sorts, it was thought that it would be of some interest to report briefly a portion of the results obtained from tabulating these cards. Previous analysis of Tippett's data have been made by Yule[2] and by Dodd.[3]

The table contains 41,600 digits printed on twenty-six pages, each page containing thirty-two columns and fifty rows; on each page the columns are assembled in groups of four. The fifty digits of each column were punched onto a card, one card for each column, thus providing 832 cards that contained all the numbers in the table. In order to check the punching the entire procedure was repeated independently on another set of cards and checked mechanically.

Table I gives the frequency of each digit and a few summary findings for runs. The runs were counted mechanically on a tabulator as for one continuous series of digits constituted as follows. The top row of the first page was made continuous with the top row of the second page, this with the top row of the third page, and so forth, until the top row of the last page, 26, was completed. The latter was then made continuous with the second row of the first page, and the process continued through the fiftieth row of the last page to complete the series.

TESTS OF RANDOMNESS

The order of the digits can be examined, of course, by a great variety of statistical tests, each testing the "null hypothesis" that the digits are arranged at "random." The various tests that can be constructed will be differently sensitive to different kinds of systematizations,[4] and by

[1] L. H. C. Tippett, *Random Sampling Numbers*, Cambridge University Press, 1927.

[2] G. U. Yule, "A Test of Tippett's Random Sampling Numbers," *Journal of the Royal Statistical Society*, Vol. 101, pp. 167–172, 1938.

[3] E. L. Dodd, "Certain Tests for Randomness Applied to Data Grouped into Small Sets," *Econometrica*, Vol. 10, pp. 249–257, July–October 1942.

[4] Joseph Berkson, "Tests of Significance Considered as Evidence," this Journal, Vol. 37, pp. 325–335, September 1942.

J. Neyman, *Lectures and Conferences on Mathematical Statistics*, Washington, The Graduate School of the United States Department of Agriculture, 1938, Chap. 3, pp. 33–48.

noting which gives unlikely values for the test P, one might be able to discern what sort of systematization was present. A number of tests that were tried are given here as examples.

1. The expected frequency for any digit and its standard deviation are given as pn and $\sqrt{pqn}$ respectively. Any observed frequency, for example, 4,264 for the digit zero, can be tested by the ratio (1) which will be normally distributed (for large n) in random samples with mean zero and unit standard deviation.

$$\frac{4{,}264-4{,}160}{\sqrt{0.1 \times 0.9 \times 41{,}600}} = 1.7, \qquad P = 0.09. \tag{1}$$

In this manner each of the digits can be tested separately to see whether there is any indication of preference or avoidance of particular digits.

2. An over-all test that will, however, not be so sensitive or revealing as to preference or avoidance of particular digits is the χ^2 frequency test which is given by (2).

$$\chi_9^2 = \frac{\sum (f - 4{,}160)^2}{4{,}160} = 11.2, \qquad P_{\chi_9^2} = 0.26. \tag{2}$$

3. *A priori* it would seem that tests based on "runs," that is, repetitions of the same digit, would be sensitive to likely departures from true randomization. It has been found by experience that it is practically impossible to maintain a random arrangement of "runs" in the presence of any kind of bias. An article by Mood[5] makes a number of such tests easily accessible.

For the number of runs of length j for digit i the mean and variance are given in (3) and the distribution for large numbers is sensibly normal.

$$\begin{aligned} E(r_{ij}) &= p_i{}^j q_i[(n - j - 1)q_i + 2] \\ \sigma_{r_{ij}}{}^2 &= p_i{}^{2j} q_i{}^2\{(n - 2j)(n - 2j - 1)q_i{}^2 + (n - 2j)q_i(1 + 5p_i) \\ &\quad + 6p_i{}^2 - [(n - j - 1)q_i + 2]^2\} \\ &\quad + p_i{}^j q_i[(n - j - 1)q_i + 2]. \end{aligned} \tag{3}$$

For instance, to test the runs of two of digit zero (Table I)

$$E(r_{02}) = (0.1)^2(0.9)[(41{,}600 - 2 - 1)(0.9) + 2] = 337.0,$$

$$\sigma_{r_{02}}{}^2 = 321.2$$

$$\frac{362 - 337}{\sqrt{321.2}} = 1.4, \qquad P = 0.08.$$

[5] A. M. Mood, "The Distribution Theory of Runs," *Annals of Mathematical Statistics*, Vol. 11, pp. 367–392, December 1940.

That is, the probability of getting 362 runs or more of two in random samples is 0.08. If there were an abnormal tendency to pairing of this digit, one would be likely to get a much smaller value for P using this test. Hence, there is little reason to suspect any abnormal pairing here. Similarly, each of the observed frequency of runs for the different digits can be tested.

TABLE I

FREQUENCY AND "RUNS" OF EACH DIGIT

i Digit	Frequency	r_{ij}, runs of length j					r_i Total runs
		1	2	3	4	5	
0	4,264	3,406	362	39	3	1	3,811
1	4,081	3,248	356	35	4	0	3,643
2	4,233	3,452	333	29	7	0	3,821
3	4,224	3,408	344	37	3	1	3,793
4	4,171	3,377	330	38	5	0	3.750
5	4,144	3,356	329	42	1	0	3,728
6	4,051	3,262	339	33	3	0	3,637
7	4,084	3,311	325	33	6	0	3,675
8	4,144	3,328	366	24	3	0	3,721
9	4,204	3,409	332	37	5	0	3,783
Mean	4,160	3,355.7	341.6	34.7	4.0	0.2	3.736.2
Expected mean	4,160	3,369.6	337.0	33.7	3.4	0.3	3,744.0

4. An over-all test for the runs of any particular digit can be made from the total runs observed for the digit. The expected value and variance are

$$E(r_i) = np_iq_i + p_i^2$$
$$\sigma_{r_i}^2 = np_i(1 - 4p_i + 6p_i^2 - 3p_i^3) + p_i^2(3 - 8p_i + 5p_i^2). \tag{4}$$

For example, to test the total runs of digit zero (Table I)

$$E(r_0) = 3{,}744, \qquad \sigma_{r_0}^2 = 2{,}733$$
$$\frac{3{,}811 - 3{,}744}{\sqrt{2{,}733}} = 1.3, \qquad P = 0.10.$$

Similarly, each of the other digits can be separately tested.

5. An over-all test can be made for the total runs for all the digits. It has been shown that the variable[6]

$$Q = \sum_1^k \sigma^{ij} x_i x_j \tag{5}$$

is asymptotically distributed according to the χ^2 law with k degrees of

[6]Formula 5 is a quadratic form and the j is not the length of run as in 3.

freedom. In (5) the x_i are defined as

$$x_i = \frac{r_i - np_iq_i}{\sqrt{n}}, \qquad i = 1, 2, \cdots, k$$

and $\|\sigma^{ij}\|$ is the inverse of

$$\sigma_{ii} = p_iq_i - 3p_i^2q_i^2$$
$$\sigma_{ij} = -p_ip_j(1 - 2p_i - 2p_j + 3p_ip_j).$$

The x_i are readily determined from the data given in Table I, with $n=41{,}600$; $p_i=0.1$; $q_i=0.9$ and $k=10$. For this particular set of data $\|\sigma^{ij}\|$ becomes

$$\sigma^{ii} = (2 - 3p)/p(1 - p)(1 - 2p)$$
$$\sigma^{ij} = (1 - 3p)/p(1 - p)(1 - 2p), \qquad i \neq j.$$

Substituting the numeric equivalents in (5), $Q=15.3$ which as a value of χ^2 for ten degrees of freedom gives $P=0.12$.

6. A quick over-all test can be made for the runs of all the digits in terms of the expectation and variance for such runs. The values are

$$E(r) = n(1 - \sum p_i^2)$$
$$\sigma_r^2 = n[\sum p_i^2 + 2\sum p_i^3 - 3(\sum p_i^2)^2]. \tag{6}$$

Thus, the test for the 37,362 runs in this series is

$$E(r) = 37{,}440, \qquad \sigma_r^2 = 3{,}744$$
$$\frac{37{,}362 - 37{,}440}{\sqrt{3{,}744}} = -1.3, \qquad P = 0.10.$$

The fact that a variety of tests have not disclosed any indication of systematization in runs does not mean that none exists. It may be that the particular tests used were not sensitive to the particular type of systematization which existed. The kind of systematization which would be introduced is likely to be related to the manner in which the table was constructed. Since the digits are grouped in the table in groups of four, the possibility that there is systematization within these groups suggests itself. Table II gives the number of runs for each digit found in the 10,400 groups of four digits in Tippett's table. The runs may be tested by the same theorems as utilized before or by simple extension of them.

TABLE II

OBSERVATIONS OF "RUNS" IN 10,400 GROUPS OF FOUR DIGITS

	Mean		Digit									
	Expected	Observed	0	1	2	3	4	5	6	7	8	9
Runs of one	3,556.8	3,547.5	3,582	3,431	3,666	3,608	3,564	3,540	3,474	3,477	3,534	3,599
Runs of two	271.4	274.7	303	291	256	276	266	266	259	273	282	275
Runs of three	18.7	19.7	24	20	17	20	25	24	17	19	14	17
Runs of four	1.0	1.0	1	2	1	1	0	0	2	1	1	1

7. As an example, to test the number of runs of two of digit zero (Table II) the expected value for a single group of four digits and the variance are given by (3)

$$E(r_{02}) = 0.0261, \qquad \sigma_{r_{02}}^2 = 0.0254.$$

For the sum of such runs in 10,400 groups the expected value and variance are

$$E(r) = 0.0261 \times 10{,}400 = 271.4$$

$$\sigma^2(r) = 0.0254 \times 10{,}400 = 264.2$$

$$\frac{303 - 271.4}{\sqrt{264.2}} = 1.9, \qquad P = 0.03.$$

The observed mean value for runs of two for all ten digits was 274.7 as compared with an expected value of 271.4. A test[7] gives

$$\frac{274.7 - 271.4}{\sqrt{\frac{271.4}{10}}} = 0.6, \qquad P = 0.27.$$

Thus there is no evidence here either of systematization. Other tests have been applied with similar results.

I wish to express my thanks to Dr. Joseph Berkson for suggesting this survey and for guidance in its conduct.

[7] The correlation existing between the variables has been disregarded.

QUALITY CONTROL APPLIED TO BUSINESS ADMINISTRATION*

By Wm. B. Rice
Plomb Tool Company, Los Angeles

The techniques of quality control, as established by Shewhart,[1] Simon,[2] and others, are being used extensively and with great success in the manufacturing problems of industry, but it is not widely recognized that these techniques can be applied to the problems of office and financial management (cf., however, footnote 7).

It has well been said that quality control is one-tenth statistics and nine-tenths engineering judgment. If, in considering the non-manufacturing problems of business administration, executive judgment is substituted for engineering, it is possible to set up a criterion by which the usefulness of the application of quality control methods can be gauged. That criterion should require the purely statistical techniques to reinforce the seasoned conclusions of experienced, expert business managers. This test is being met successfully in several experiments conducted by the author, at the same time uncovering significant facts which executives can use in the direction of efficiency and profit.

Under the pressure of war-time conditions, many firms with contracts negotiated at fixed prices are finding that overtime wages paid to the office force become a noticeable factor in rising costs, a factor for which it is difficult to get compensation through the process of re-negotiation. The *preliminary* data used by the author in establishing economical limits for office overtime pay consisted in the percentage of pay for overtime work to pay for regular work, by weeks, for six departments, from January through June 1942.

If we think of each department as a different machine operator, and take for each "production sample" the figures for the four (or five) weeks in each month, we get six points for each of the six months, a total of thirty-six samples in all. The preliminary data were set up in tabular form as shown in the table. Beginning with July, the data were improved by eliminating a non-homogeneous element—the salaries of supervisory personnel whose compensation was on a monthly basis and who did not get overtime pay. Calculations made from the new figures

* A report presented at an Engineering, Science and Management War Training Course on Quality Control held at Los Angeles in the Fall of 1942 under the auspices of Stanford University, and sponsored by the U. S. Office of Education. The original presentation has been revised to carry it through December 1942.

Credit and thanks are due to Professor Holbrook Working, of Stanford University, and to Dr. W. Edwards Deming for their inspiring help in preparing this paper for publication.

[1] W. A. Shewhart, *Economic Control of Quality of Manufactured Product*, D. Van Nostrand, 1931.

[2] Leslie E. Simon, *An Engineers' Manual of Statistical Methods*, John Wiley & Sons, 1941.

for July, August, and September, raised the overall average from 13.4 to 20.3 per cent, reflecting a truer picture of the facts; and decreased the average range, as was to be expected, from 12.4 to 10.6 per cent.

RELATION OF OVERTIME PAY TO PAY FOR REGULAR WORK, BY WEEKS AND DEPARTMENTS

(*Per cent*)

	Weeks						
	1st	2nd	3rd	4th	5th	Average ($\overline{X}$)	Range (R)
January 1942							
Department 1	20.4	16.6	11.4	12.2	—	15.2	9.0
2	28.5	2.4	3.7	7.4	—	10.5	26.1
3	25.0	11.4	10.4	9.7	—	14.1	15.3
4	9.5	16.0	8.2	11.2	—	11.2	7.8
5	22.3	27.0	24.5	36.0	—	27.4	13.7
6	7.5	10.4	7.9	12.4	—	9.6	4.9
February 1942							
Department 1	17.1	16.1	7.8	9.9	—	12.7	9.3
2	5.1	5.6	4.2	7.8	—	5.7	3.6
3	11.6	10.2	7.6	16.3	—	11.4	8.7
4	12.2	15.5	11.6	12.0	—	12.8	3.9
5	21.2	18.3	34.3	48.3	—	30.5	30.0
6	15.9	15.2	11.3	15.1	—	14.4	4.6
			Etc.				

Using the new data for analyzing past experience, and the notation of the American Standards Association 1942 pamphlet[3] the central lines and action limits were computed as shown below.[4,5]

$$\overline{\overline{X}} = 20.3 \text{ per cent}$$
$$\overline{R} = 10.6 \text{ per cent for samples of 4}$$
$$\overline{\overline{X}} \pm A_2\overline{R} = 28.0 \text{ per cent and } 12.6 \text{ per cent}$$
$$D_3\overline{R} = 0$$
$$D_4\overline{R} = 24.2 \text{ per cent}$$

For several months in the first half of the year department 5 was "out of control" on both average and range (see chart). In June, an alert office manager caught the excessive overtime of department 5, hired more help, and promptly thereafter brought it back within normal limits. Needless to say, if the control chart technique had been avail-

[3] American Standards Association, *Control Chart Method of Controlling Quality During Production*, Z1.3-1942 (American Standards Association, 29 West 39 Street, New York), p. 39.

[4] The weeks containing July 4th and Labor Day were omitted because they were abnormal. Consequently, all three months contained four weeks.

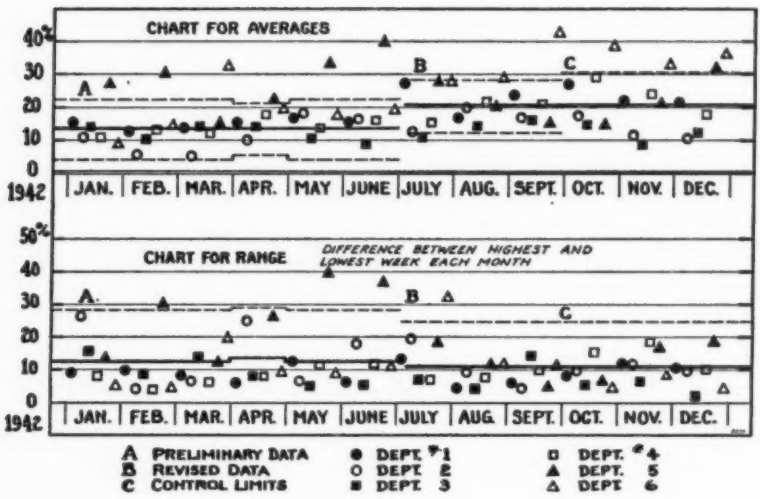
[5] The quantities A_2, D_3, and D_4 are needed for computing the control limits on the charts. Their numerical values are given in tables; e.g. in Simon, *loc. cit.*, pp. 204 ff., and in the appendix to the American Standards publication listed in footnote 3.

able, indications of trouble would have been given much earlier, with a worth-while saving in labor costs.

More recently department 6 has gone out of control; but the difficulty of finding enough trained men for a specialized type of work under present labor market conditions is hindering efforts to remedy the situation. This, however, is not so much a statistical as a personnel problem.

CONTROL CHART FOR OFFICE OVERTIME
PER CENT OF OVERTIME PAY TO REGULAR PAY

(Each point is the average for one month)



Management's opinion as to the most economical amount of overtime allowable varies according to circumstances. Normally, the standard might be set at no overtime, i.e., $\overline{X}=0$. Under war conditions, when volume of production is paramount and workers are hard to get, a certain amount of overtime work is often necessary. But the decision is difficult as to what limits should be set. They should certainly not be determined without operating facts to go on, just as engineering specifications for a manufactured article should not be drawn up without a knowledge of what limits inhere in the process.

In both factory and office quality control methods develop a basis upon which to base executive decisions.

How well such an application to overtime wages meets a commonsense criterion is indicated by the upper limits for averages derived from the data for July, August, and September. Now it happens that

this limit, 28.0 per cent ($\overline{\overline{X}}+A_2\overline{R}$ above), is close to 30 per cent, which represents 20 per cent overtime in hours worked, or one full extra day in a normal five-day work week. In this case the management considered it reasonable to argue that if any department had enough work to keep its staff busy for an extra day, economy would be served by hiring sufficient additional help to do the work in five days, thus avoiding the necessity of paying overtime premiums regularly.[8] With these thoughts in mind, the control limits for future data were set up as follows:

For averages:	Upper limit, 30 per cent
	Lower limit, 0 per cent
For ranges:	Upper limit, 24.2 per cent
	Lower limit, 0 per cent

These control limits were applied to the last quarter of the year, i.e., October to December 1942. The weeks containing holidays were omitted from the calculations because, with a full working force present on those days, overtime payments were necessarily abnormal. Effectiveness of quality control is indicated by the fact that with the exception of department 6 and one point for department 5, all overtime was kept within the agreed-upon limits. Seasonal factors accounted for department 5's unavoidable lapse in December.

In this connection the range chart is particularly important. Showing as it does the difference between the highest and lowest week in amount of overtime paid for during the month, the range gives a measure of the skill and care with which work-flow through each department is planned and supervised. Extreme irregularity in work-flow shows up as an out-of-control point on the range chart, and can often be used to disprove a department head's alibi of "too much work" when he is confronted with the fact of excessive overtime paid to his employees.

The author's observation is that, in applying quality control to general management, the behavior of the range or standard deviation chart will reveal more valuable information than the chart for averages.

Office overtime analysis is only one example of the application of quality control to the interests of general management. The field is fertile but largely unexplored. For instance, quality control can be used to set norms and action limits for clerical work, both by individuals and by groups where, as in stenography or billing, quantitative standards can be established. An application toward the improvement of clerical work, and toward reduction in costs, has been made and re-

[8] With the 48-hour, six-day week established officially in many sections of the country, these conclusions may need to be reconsidered.

ported by Dr. Deming and his staff at the Bureau of the Census.[7] Evaluation of sales and territories, by salesmen and by lines, can be undertaken. Budgetary and cost accounting records are illuminated. Along some of these lines work now being done by the author indicates that, as in manufacturing, such studies do determine economic action limits, thereby lessening the heavy burdens of executive responsibility.

To the many friends of quality control who are personally convinced of its value but who have to sell their superiors on the idea, its application to specific problems in business administration offers an inexpensive and powerful psychological approach. It requires no costly experimental interruptions of production; treads on few toes; invites fewer organizational upheavals; and by proving its worth in one field, opens the way logically to many others.

[7] W. Edwards Deming and Leon Geoffrey, "On Sample Inspection in the Processing of Census Returns," this JOURNAL, Vol. 36, 1941, pp. 351–360.

FAMILY VERSUS INDIVIDUAL IN MEASUREMENT OF AUDIENCES

By Alfred Politz

Advertising research owes its place entirely to the requirements of advertising itself. Thus the value of a particular research tool depends entirely and only on its usefulness to advertisers.

Since the best of advertising has no selling power until it is presented, one of the most fundamental problems to advertisers, publishers and radio men is to know how many people read a magazine or listen to a radio program. And it is in solving this basic problem that a number of research endeavors have adopted a method which defeats their purpose.

It came about this way. Since most people live together in groups as families, there exists the idea of the family as a social unit. But this idea goes a step further as far as advertising is concerned. Many or most articles which the members of the family consume are not bought by all members separately, but only by one representative, for example, the housewife. Thus the family assumes the appearance not only of a social unit, but also of a single consuming unit.

In view of this, it seemed to research men and publishers, and justly so, that it was inadequate to express the readership of an advertising medium in terms of total individuals. It seemed that there should be a unit of measurement which would better parallel this structure of consumption. So the "family base" was chosen. No specific studies using the family as the base need to be mentioned, since it is generally known that many, if not most, readership reports employ this method.

Assume that there is a statistical universe of 24 families. For the sake of simplicity, each family consists of four members: Husband, housewife, son, and daughter.

In this universe there are two newspapers, A and B. Expressed in terms of individuals, each newspaper has 48 readers. So in terms of individuals they are equal.

But the distribution of the individuals differs in the following way:

> A is read by 2 persons in each of
> 24 different families (housewife and daughter).
>
> B is read by 4 persons in each of
> 12 different families.

Which paper is a better medium for tomato juice?

According to certain research conclusions, A is better, because it goes into twice as many families as B.

The idea is that tomato juice is purchased by the housewife. If the ad is read by her, the purpose of the ad is practically fulfilled. Any additional reading by husband, daughter, or son is of only secondary value. These persons may speak to the housewife about tomato juice and increase her willingness to buy a certain brand, but it is obvious that such influence upon the final purchase is smaller than that of the housewife herself.

A and B are equal, inasmuch as each is read by 48 persons.

Newspaper A, however, which goes into 24 families, influences 24 housewives. While newspaper B, which goes into 12 families, influences only 12 housewives. And so, on the "family base," A is better than B.

Thus there are many research men and advertisers who believe that the family, instead of the individual, is the right unit for describing the readership of newspapers and magazines and the audience of radio programs. But the whole family approach is based on a fallacy.

It is true that, in this particular example, the number of families comes closer to the description of the relative values of the newspapers than does the total number of individuals. But that is more or less accidental and does not warrant the use of "the family" as a yardstick. *For what really was used as the gauge was the number of housewives.*

The family is only a condition of social life which gives the housewife (an individual) a greater importance for buying tomato juice than the son (another individual). Consequently an advertiser of tomato juice is more interested in the housewife than in the son.

There are other conditions, also, which determine the value of certain individuals.

Some people's habit of living in apartment houses makes the superintendents, who also live there, more important to coal dealers than the ordinary tenants. Southerners live in a warmer climate than the Northerners, and so are more important to manufacturers of light clothing.

There is a direct and simple device in every statistical picture which takes care of situations in which one type of individual is distinguished from other types. That is the breakdown of the total figure into specific ones.

Suppose, for instance, that the total figures of readers are known for different magazines, and that Southerners are more important for the given purpose than Easterners, Westerners, and Northerners. Then, from the total figures, the number of Southerners within the readership of each magazine simply must be sorted out.

There is a universe of people living in apartment houses. Two newspapers are distributed among them, and both newspapers are equal in

the total number of individual readers. But the particular advertising problem is concerned with coal. Therefore the total number of readers is not sufficient. The figure must be broken down into "superintendents" and "tenants." If it appears that one paper is read by more superintendents than the other, it therefore has a better chance to sell coal.

If lipstick is the product, the readership figures are broken down into females and males. For Cadillac cars, they are broken down into well-to-do people and not well-to-do.

For tomato juice, the figures must be broken down into housewives and others.

If paper A with 48 readers reaches 24 housewives, and B with its 48 readers reaches only 12 housewives, then A is better than B. It is better because it reaches more housewives and not because it reaches more families.

It is possible that the greater number of housewives may coincide with a greater number of families. And in the first "tomato juice" example, it actually did.

Any different example, however, is equally justified.

Suppose, as in the first example, that paper A has 48 readers in 24 different families. And it is always read by two family members—in this case, however, son and daughter, instead of housewife and daughter.

Paper B is read entirely as before. It has 48 readers in 12 families, and it is always read by all four family members, husband, housewife, son, and daughter.

If, for the sake of tomato juice, the number of *families* indicates the value of a newspaper, then A is still better than B, for A still reaches twice as many families.

But looking at the kind of *individuals* it reaches—there is not a single housewife involved. Although there are twice as many families, it is read only by sons and daughters. So it cannot have twice as great a value; quite on the contrary, it may have no value at all.

Therefore as was said before, if newspaper A has more value for tomato juice, it cannot be on account of the greater number of *families*, but only on account of the greater number of *housewives* it reaches. And it is a mere accident if both figures happen to be in agreement. (As a matter of fact, they usually are in disagreement, for the very reason that a family is a collective of heterogeneous elements in which all ages and sexes are mixed.)

The family, itself, has neither a sex nor an age nor an occupation in which an advertiser might be interested. It cannot read and it cannot buy, the two activities which are most important to an advertiser. And

whenever an attempt is made to use such a unit as a yardstick in measuring activities which it does not perform, but which are performed rather by one or two or three or more of its members, the purpose of the yardstick is defeated.

If surveys using the family base speak of 1000 families which are reached by a newspaper or magazine, they do not define what they mean. Do two members of each family read the newspaper? Or does one member of each family read the paper? Or is a family counted if at least one member reads the paper? In practice the "at least one" definition is used.

"At least one" means: One, or two, or three, or more. There is the first variable; it is not known how many people read the newspaper.

But there is still another variable. If a paper reaches 1000 families—and that means it reaches either 1000 or 3000 or more people—who are these people? They are not housewives, for if they were, they certainly would have been called so. Nor, for the same reason, are they husbands. They are, in fact, not described at all.

In a short sentence it looks like this: A newspaper which "reaches 1000 families" is a newspaper which is read by 1000 or 3000 or more people, who are either housewives or husbands or someone else.

The variables in this definition render the "family" useless as a yardstick. There is a possibility, however, of making the "family" useful as a yardstick by attaching to it a sufficiently rigid definition. That would be: A family is reached by a newspaper or radio program if each of its members (or adults) is reached.

On the basis of this reasonable definition, however, what must the research man do with a copy which is not read by each member of a family, but by only one or two or three members of a family which consists of four? In accordance with the last definition, such a copy does not reach the family.

The number of families reached would be correct and clear in itself, but very small. Only a minority of copies reaches every member of a family. So by this definition the majority of copies is neglected. In other words, the figure so obtained does not say anything about the fate of the copies, but only about the fate of some of the copies. It does not tell about readers of the paper; it tells only about certain readers, namely, those who have close relatives reading the same paper. The definition is technically clear but irrelevant to the purpose. This is so obvious that it automatically explains why this definition never was used.

It must be remembered that there is a priori so much clarity and simplicity in expressing the value of a newspaper or magazine in terms

of individuals, that the inventors of the family base needed an enticing reason for becoming involved in it. The enticing reason was the fact that equal numbers of individual readers do not necessarily express equal values of newspapers. Paper A reaches 48 readers in 24 families (husband and housewife are readers). Paper B reaches 48 readers in 12 families (husband, wife, son, and daughter are readers). For selling tomato juice, newspaper A is better than B. "While having the same total of individual readers, it reaches twice as many families." Now this statement is only true if a family can be constituted by two of its members as well as by four.

Thus it develops that the "family" could offer its delusive service as a yardstick only *because* of its hidden flexibility.

It is understood that no one tried to aim at flexibility. The aim was to use a unit which would parallel in audience measurement the known facts of family consumption. Unfortunately "the family" was chosen as the unit, without sufficient analysis. It provided the chance to use comfortable terminology and gave the feeling of scientific rigor. Yet by its very nature it led to erroneous conclusions.

On the other hand it requires only a very small sacrifice of preliminary comfort to obtain and enjoy an accurate method of measuring reading and listening audiences. It is the individual who listens and reads; it is a certain type of individual whose reading and listening are more or less important for a particular purpose. And it is the statistical sorting out of these types which tells the complete story about the value of a radio program or a periodical.

THE COST OF LIVING IN CITIES IN RELATION TO THEIR SIZE AND LATITUDE

By E. L. Thorndike and Ella Woodyard

It is repeatedly asserted or assumed that the cost of living is much greater in large cities than in small, and in northern cities than in southern. Both statements are false. Except for rent the differences in retail prices between cities of 30,000 to 44,000 inhabitants and cities of 300,000 to 500,000 inhabitants are near zero and the difference in rent for a house or apartment of equal quality is less than twenty per cent. The differences in relation to latitude are almost nil.

We examined the advertisements appearing September 13 to 18, 1937, in the daily newspapers of the cities (265 in number) of over 30,000 population and under 500,000 (in 1930) which had daily newspapers.[1] We recorded carefully the prices found in all advertisements for any of the 28 items listed below, except that care was taken not to include more than once an apartment or house advertised on more than one day or by more than one agent, and that if more than twenty of any one size were advertised, twenty were chosen at random.

1. Monthly rent, apartment of 3 rooms, unfurnished
2. Monthly rent, apartment of 4 rooms, unfurnished
3. Monthly rent, apartment of 5 rooms, unfurnished
4. Monthly rent, apartment of 6 rooms, unfurnished
5. Monthly rent, house of 4 rooms, unfurnished
6. Monthly rent, house of 5 rooms, unfurnished
7. Monthly rent, house of 6 rooms, unfurnished
8. Monthly rent, house of 7 rooms, unfurnished
9. Monthly rent, house of 8 rooms, unfurnished
10. Sale price of house of 4 rooms, unfurnished
11. Sale price of house of 5 rooms, unfurnished
12. Sale price of house of 6 rooms, unfurnished
13. Sale price of house of 7 rooms, unfurnished
14. Sale price of house of 8 rooms, unfurnished
15. Sale price of house of 9 rooms, unfurnished
16. Permanent wave, lowest of the advertised prices
17. Permanent wave, median of the advertised prices
18. Permanent wave, highest of the advertised prices
19. Movies, average of lowest price of evening admission, and of other than evening admission if such is advertised
20. Movies, average of median price of evening admission, and of other than evening admission if such is advertised
21. Milk, using all firms advertising[2]

[1] Except that three resort cities (Atlantic City, Miami, and St. Petersburg) were excluded and Milwaukee was used instead of Washington.

[2] Milk prices appeared in too few cities (only 18) to be used.

22. Butter, using all firms advertising
23. Eggs, using all firms advertising
24. Chicken, using all firms advertising
25. Fowl, using all firms advertising
26. String beans, using all firms advertising
27. Potatoes, using all firms advertising
28. Cabbage, using all firms advertising

Expressing all prices as percentages of the median price for the thing in question in cities of 55,000 to 65,000, we have the facts of Table I.

TABLE I

THE RELATION OF CERTAIN ADVERTISED PRICES TO SIZE OF CITY IN 265 CITIES

Size of city	Prices expressed as percentages of the median price for cities of 55,000 to 64,999 inhabitants							Average
	Chicken	Fowl	Vegetables	Eggs	Butter	Movies	Permanent waves	
300,000 to 499,999	85½	94	100	103	104	100	106	98.9
200,000 to 299,999	94	100	110½	89	103	102	90	98.4
115,000 to 199,999	104	100	100	96	103	101½	99	100.5
95,000 to 114,999	96	104	100	98	100	91½	100	98.5
75,000 to 94,999	96½	93	105½	104	100	100	100	99.9
55,000 to 74,999	100	100	107	98	100	100	99½	100.6
45,000 to 54,999	100	100	107	96	101	103	107	102.0
35,000 to 44,999	100	100	107	104	100	85	102	99.7
30,000 to 34,999	94	96	100	96	100	100	97	97.6

We did not make similar records and computations for the retail prices of staples like flour, sugar, gasoline, trade-marked goods, drugs, books, cigars or cigarettes, because these should be less sensitive to size of city than the commodities of Table I. But a rough inspection of prices for some of them shows no evidence that they are cheaper in the smaller cities.

Expressing the median rental of an apartment or house of specified size in each group of cities as a percentage of the median price in cities of 55,000 to 65,000, we have nine sets of percentages for apartments of 3, 4, 5, and 6 rooms and houses of 4, 5, 6, 7, and 8 rooms. We also have six similar sets of percentages for sales prices of houses of 4, 5, 6, 7, 8, and 9 rooms. The averages, when all fifteen sets are given equal weight, are as shown in Table II. There is a clear rise of about 10 per cent from cities under 55,000 to cities of 55,000 to 114,999, and a less clear rise of about 5 per cent from cities of 55,000 to 114,999 to cities of 115,000 or over. Any reasonable system of weights gives closely similar results. In

so far as the apartment or house of the same number of rooms tends to be of better quality in the larger cities, the figures of Table II exaggerate the differences in rental or purchase for equally good accommodations.

TABLE II

THE RELATION OF ADVERTISED PRICES OF APARTMENTS AND HOUSES FOR RENT AND FOR SALE TO SIZE OF CITY IN 265 CITIES

Size of city	Average of fifteen sets of percentages for rents of apartments of 3, 4, 5, and 6 rooms, and houses of 4, 5, 6, 7, and 8 rooms, and of houses for sale of 4, 5, 6, 7, 8, and 9 rooms
300,000 to 499,999	102½
200,000 to 299,999	98
115,000 to 199,999	98
95,000 to 114,999	91
75,000 to 94,999	95
55,000 to 74,999	98½
45,000 to 54,999	86
35,000 to 44,999	86
30,000 to 34,999	84½

The higher cost of lodging in the larger cities is not a consequence of higher prices for building materials. The data reported by the U. S. Bureau of Standards on the prices of brick, sand, stone, cement, lime, boards, 2×4's, gypsum board, nails, tar-paper, glass, linseed oil and white lead, are in fact higher for cities under 100,000 population than for those from 100,000 to 200,000, and higher for the latter than for cities of over 200,000. Using percentages of a base-price for each commodity, the averages of the thirteen percentages are as shown in Table III.

TABLE III

PRICES OF BUILDING MATERIALS IN RELATION TO SIZE OF CITY

Size of city	Number of cities studied by the Bureau of Standards	Average of 13 percentages for brick, sand, etc.
Over 1,000,000	4	92.9
500,000 to 999,999	6	91.7
300,000 to 499,999	8	92.3
200,000 to 299,999	9	96.5
100,000 to 199,999	16	100.2
30,000 to 99,999	11	102.6
Under 30,000	3	114.9

A critical study of the costs of lodging, foods, carfares, electric current and other commodities and services in various cities reported by the U. S. Bureau of Labor Statistics and the National Industrial Conference Board reveals that the discrepancies between their findings concerning

size of city and ours are not large and may be explainable by differences in quality.

The advertised prices in the 265 cities show no clear differences in relation to latitude, either for foods, movies, etc., or for rent. The medians of eleven percentages summarizing both for the groups at 44° or more, 43°, 42°, 41°, 40°, 39°, 38°, 36° and 37°, 33° to 35°, and south of 33° are, in order, 96, 91, 100, 100, 102, 100, 98, 102, 93, and 88.

In the prices for domestic use of electricity reported for 51 cities by the Bureau of Labor Statistics there is little or no correspondence with latitude. The percentages for 12 cities north of 42°, 13 cities at 40° or 41°, 13 at 37° to 39°, and 13 south of 37°, are in order 102, 104, 90, and 105. The prices of building materials are a little *lower* in the most northerly cities.

EXPERIENCE WITH TESTS OF SIGNIFICANCE: A REPLY TO PROFESSOR R. A. FISHER

By Joseph Berkson, M.D.
*Section of Biometry and Medical Statistics, Mayo Clinic**

Professor R. A. Fisher in his note on my article "Tests of Significance Considered as Evidence"[1] finds it pertinent to say, "It is not my purpose to make Dr. Berkson seem ridiculous, nor, of course, to prevent him from providing innocent amusement," and to make some biographical comments of a similar intention. This is no time to begin calling names across the sea, and however strongly I may differ with Professor Fisher in regard to scientific questions, I shall confine my differences to the subject matter discussed.

My present attitude toward statistical tests of significance is the result of more than fifteen years of experience in their application to a large variety of data. I have no desire to undermine them unwarrantably. On the contrary, since my profession is that of consulting statistician, I should like to possess a tool in the use of which I was a specialist and which could serve well those who engaged my services. But if I have found these tests misleading in certain applications, it is appropriate that I expose this fact in statistical journals for the view of my colleagues. I may be wrong, but I am not such a fool that I cannot repeat as well as anyone who has ever memorized a textbook, the argument about the null hypothesis and the small P disproving it. At the beginning of my career I quoted the orthodox view. But a consulting statistician in a practical institution finds himself under the necessity of reconciling what *he says* to those who employ him with what *they find.* Time and time again I came upon actual situations in which (1) it was patent to all competent people who viewed the data, that the null hypothesis, on the evidence, *was not* true, but a small P did not result, (2) a small P did result, but it was obvious to all who examined the data, that on the evidence the null hypothesis *was* substantially true, or (3) when the null hypothesis was tested by several tests, each put forward with equally good authority, the P's were considerably *different.*

Reflection on this situation led me to the conclusion that the argument by which tests of significance are usually supported is basically unsound. The fallacy lies fundamentally in the argument's not con-

* At present on leave, Lt. Col., Medical Corps, Army of the United States, Chief, Statistical Division, Office of the Air Surgeon.

[1] Joseph Berkson, "Tests of Significance Considered as Evidence," this Journal 37: 325–335 (September) 1942. R. A. Fisher, "Note on Dr. Berkson's Criticism of Tests of Significance," this Journal, 38: 103–104 (March) 1943.

sidering the *alternative* to the null hypothesis. I have found that the quite numerous instances of false or paradoxical conclusions of my experience are resolved, if the view is adopted that what the result obtained from a test of significance does is not to "disprove the null hypothesis tested," nor even to give a basis for "rejecting" it but rather that the finding affords evidence *in favor* of something—in favor of some *alternative*. In ordinary rational procedures and in scientific inquiry too, one posits explicitly or implicitly a set of alternatives, A and not A: for instance, the patient has syphilis or he does not have syphilis; there is no difference in means or there is some difference; the distribution is mesokurtic or the distribution is heterokurtic; the regression is straight or the regression is curved. When the P of an appropriate t test is small, according to my view, one does not "disprove" that there is no difference in means; one has affirmative evidence in favor of the alternative, that is, that there is such a difference. I say this constitutes such evidence, because in the circumstances, a small P would occur frequently (relative to the level of significance) if there did exist a difference between the means.

The small P's of tests are evidence for any alternative in which they occur frequently. Good tests are ones in which (1) when a small P occurs it is evidence for an alternative in one particular pair of contraries and (2) when this alternative is true, a small P will occur very frequently. Such a test might be called highly specific and sensitive. An example is the t test appropriately performed with ample numbers for a difference between the means of two normal distributions. When with such a test a small P results, it points to nothing except that there is a difference in means; if there is a substantial difference in means, such a P is likely to result.[2] Tests which Neyman and Egon Pearson call "universally most powerful" tend to be specific in this sense. When with such specific tests one has sufficient numbers, they become sensitive tests; in the terminology of Neyman and Pearson they become "powerful." I have suggested the generalization of which I feel only fairly sure, that highly specific tests are always equivalently statable in terms of an estimate and its confidence limits, and that it is better to state them in that way. There are being widely used in statistics today tests that are not good tests, in the sense mentioned, and these are the ones that I have been questioning. Examples are (1) the frequency chi-square test and (2) many applications of analysis of variance.

If the view is adhered to that what the small P of a test of signifi-

[2] J. Neyman and B. Tokarska, "Errors of the Second Kind in Testing 'Students,' Hypothesis, this JOURNAL 31: 318–326 (June) 1936.

cance does is to "disprove the null hypothesis tested" then not only will these latter tests be poor tests in being nonspecific or insensitive or both; they may be positively misleading. The case I cited of Professor Fisher's application of the technic of analysis of variance as a test for linearity is a case in point. This example was not brought out in isolation to point up anyone's individual blunder, but was part of a continued discussion of the pitfalls into which the pursuit of an erroneous conception of tests of significance could lead one. It was only one of many examples that I have published, mostly from my own experience. Professor Fisher applied[3] what he titled a "test of straightness of regression line" to some data of Hersh,[4] of eye facet number of Drosophila in relation to temperature, obtained a small P and concluded that the regression is nonlinear. He insists in his recent note that this is so, because this, he says, is the objective way of determining that it is not linear. Now as a naive desk neighbor of mine said when I read him the results, "If it isn't linear, what is it?" A plot of the points and a least square fit did not indicate any curvature that I or anyone else to whom I exhibited the graph was impressed with. If I were to predict facet number from a regression on temperature for these data, I should use a linear regression. On these grounds I wrote: "On inspection it appears as straight a line as one can expect to find in biological material." Professor Fisher thinks this judgment of the lack of curvature, instead of the reverse judgment based on the small P of his test, is obtuse and authoritarian. But I presented the evidence to the reader and I gave another explanation of the small P.

I do not adhere to the doctrine that a small P disproves the "null hypothesis tested," in this case that the regression is linear. I have attempted to demonstrate in the article mentioned that, while this principle appears plausible, it is actually illogical and at variance with the usual mode of scientific demonstration. I do not think that a logician would support the proposition that "If an event known to occur rarely in hypothesis H, occurs, hypothesis H is disproved," and no validating experiments have been brought forward to support it either. In practical fields of investigation in which a test, T, is utilized, as for instance when a serologic reaction is used for the diagnosis of syphilis, validation of the test involves the determination of the frequency of T in (a) normal persons (null hypothesis) and (b) syphilitic patients (alternative). No one in this field would support such a test by demonstrating merely

[3] R. A. Fisher, *Statistical Methods for Research Workers*. Ed. 7, Edinburgh, Oliver and Boyd, 1938, 356 pp.

[4] A. H. Hersh, "The Effect of Temperature upon the Heterozygotes in the Bar Series of Drosophila," *Journal of Experimental Zoology*, 39: 55–71 (February) 1924.

that if the population is normal, T occurs only rarely. Such an argument would appear absurd. What is wanted also, and indeed what is wanted mainly, is information as to how *often* T occurs if the population is syphilitic. It is this that determines the usefulness of the test for its purpose, the diagnosis of syphilis.

Why should the validity of statistical tests be immune to such critical appraisal as is required as a matter of course in any other field of application of tests? Analogous validation for statistical tests of significance would involve determining, first in what *alternatives*, and then how *often* the critical value of the test function occurs. Before the work of Neyman and Egon Pearson, no such inquiry was made for *any* statistical test of significance. Even now it has not yet been made for the tests that I have been questioning, despite raucous and repeated challenges from me to the mathematicians; for instance, it has not been made for all phases of the test of linearity of Fisher which we are discussing.[5] Examples of questions which would be answered by such an inquiry are: What is the probability of obtaining a value of $P \leqq 0.05$ with this test if (a) the abscissal variate x is not subject to variation, the ordinate variate y normally distributed with uniform variability σ_y, and the means of $\bar{y}_x$ follow some given nonlinear regression, (b) the regression is linear, the abscissal variate not subject to variation but the variability of y, σ_y is a given function, not constant, (c) the regression of means is linear, σ_y constant, but x not constant but subject to normal variation with σ_x?

I say the small P is evidence for any alternative hypothesis in which it will occur frequently (relative to the level of significance). In the instance in hand I suggested three[6] specifically and others generally: (1) curvilinearity as against linearity, (2) heteroscedasticity as against

[5] P. C. Tang, "The Power Function of the Analysis of Variance Tests," *Statistical Research Memoirs*. Edited by J. Neyman and E. S. Pearson, Department of Statistics, University College, London, December, 1938.

[6] At several points Professor Fisher achieves miscomprehension only with great effort. He says: "What Dr. Berkson fails to realize is that the judgment, from inspection, that the line appears as straight as one ought to expect, would have given no aid whatever towards discovering the cause of any real anomaly, whatever the cause might be, because that judgment in effect denies the evidence that any real anomaly exists." The fact is that it was I, not Professor Fisher, who suggested the multiplicity of possible meanings of the small P in this application. I said: " . . . a small P is evidence in favor of curvilinearity . . . , also . . . in favor of linearity plus heteroscedasticity . . . the temperature . . . subject to fluctuation . . . and . . . other conditions which . . . produce a small P relatively frequently . . . Which of these shall be taken to have been demonstrated . . . will have to be determined by other evidence, possibly other statistical tests." It is hard to imagine how I could have been plainer on this point. The possibility of a "remarkable discontinuity . . . in the neighborhood of 24°C.," not mentioned in Professor Fisher's textbook illustration, which is what I was commenting on, is provided for in the last part of my statement. We were discussing, however, if it requires recollection, not a test for *anything at all*, not just to see whether *something* was the matter. We were talking about a test for "straightness of regression line," that is a device by which we were to discern whether the regression was linear or curvilinear, and a specific application of it in which the regression had been found definitely to be nonlinear.

homoscedasticity, (3) variable temperature as against constant temperature, and (4) other conditions which will produce a small P. I gave as my own preference, in view of the fact that (to me) there was no confirmation of the nonlinearity, that the temperature was fluctuating. Examination of the figure suggests this explanation as especially likely if the temperature was variable at what was recorded as 23 or 25°C. or both. Professor Fisher says that, if I had read Hersh's article from which the data were obtained, I should not have committed the blunder which I made. I had read the article and it amply confirms my suggestions of possible alternative causes of the small P. There is definite evidence of heteroscedasticity in table 6 of the article, and as for fluctuation of temperature, here are Hersh's words (my italics):

> Errors in this study may arise from fluctuating temperatures. The fluctuations were not the same in all cases; 15°, 17°, 25°, and 27° showed a variation of ±0.5°. 23° varied as much as 3°. The record of 29° showed a variation from 28.0° to 29.5°. Temperature readings taken of the food by a thermometer graded in tenths of degrees, showed a total variation of 1.4° in the 31° aquarium, and of 1.3° in the 21° aquarium. The 19° aquarium was more unsatisfactory and *varied sometimes as much as 5°*. In every case, however, an attempt was made to have the proper temperature during the facet reaction period, though *this may not have been accomplished* because the period may not be the same for the heterozygotes as for the homozygotes for which it has been determined. Nominal temperatures in accompanying tables and figures are to be considered with these remarks in mind.

EDWARD LEWIS DODD, 1875–1943

THE SUDDEN DEATH of Edward Lewis Dodd on January 9, 1943, at his home in Austin, Texas, following a heart attack, brought to a close a distinguished career. He had long been a member of the American Statistical Association and had just been elected a Fellow of the Association, though the news of his election did not reach Austin until shortly after his death.

Dr. Dodd was born in Cleveland, Ohio, November 1, 1875, the son of Heber and Emma Melvina (Drake) Dodd. He was educated in the Cleveland public schools; in Western Reserve University, from which he received the degrees of B.A. in 1897 and M.A. in 1901; and in Yale University, from which he received the degrees of M.A. in 1902 and Ph.D. in 1904. After serving as instructor in mathematics for two years at the State University of Iowa and one year at the University of Illinois, he came to The University of Texas as instructor in pure mathematics in 1907, and rose through the successive ranks of adjunct professor (1914), associate professor (1917), and professor (1923) of actuarial mathematics. In the year 1922–23, he taught at Williams College as visiting assistant professor of mathematics. On June 15, 1910, Dr. Dodd was married to Miss Alice Laidman of Austin, Texas. His wife survives him.

Dr. Dodd was a member of the honorary societies of Phi Beta Kappa and Sigma Xi and of various professional societies in addition to the American Statistical Association, including the American Mathematical Society, the Mathematical Association of America, the Econometric Society, and the Institute of Mathematical Statistics (of which he was a Fellow). He was a councilor of the Mathematical Association of America in 1920, and Associate Editor of the *American Mathematical Monthly* from 1918 to 1923.

His main fields of work were probability, statistics, and actuarial mathematics. During its early days he was a member of the Actuarial Advisory Committee of the Teachers Insurance and Annuity Association of America. He also gave much assistance in connection with the establishment of the Teachers' Retirement System of Texas.

Of the many honors which came to Dr. Dodd in recognition of his attainments in his chosen field, two will be mentioned here. He was chosen as the official representative of the Mathematical Association of America at the Seventh American Scientific Congress held at Mexico City, September 8–17, 1935. He was one of the two Americans who were invited to participate in the International Congress on the Theory of

Probability held at Geneva, Switzerland, October 11–15, 1937, and was the only American who actually did take part in it.

Dr. Dodd was an inspiring teacher, a distinguished scholar, and a master of productive research. A list of his publications, in addition to numerous reviews, abstracts, and minor papers, includes some forty items. But to those who knew him best, he will also be remembered for his sincere friendliness, his unfailing courtesy, and genuine kindliness, and his being always the Christian gentleman.

C. D. SIMMONS

The University of Texas

THE STUDY OF CYCLES

THE FOUNDATION for the Study of Cycles—a non-profit organization created to foster, promote, and conduct scientific research in respect to rhythmic and periodic fluctuations in any branch of science—announces the offering of a medal to the person who, during 1943, publishes the book or paper in this field that in the opinion of the judges is the most outstanding.

The judges are Dr. C. G. Abbot, Solar Physicist and Secretary of the Smithsonian Institution; Dr. Harold E. Anthony, Curator of Mammals and Dean of Scientific Staff, American Museum of Natural History; Professor W. C. Mitchell, Professor of Economics, Columbia University, and Director of the National Bureau of Economic Research; Professor V. C. Wynne-Edwards, Professor of Zoology, McGill University; and Professor Ellsworth Huntington, Professor of Geography and Climatology, Yale University (Chairman).

It may be in point to recite some of the reasons for assigning high importance to research in cyclic phenomena. First, in dealing with rhythm and periodicity the scientist is at the very heart of predictability. The power to predict accurately is the acid test of the degree of precision reached in any science. Medical science, for example, would become vastly more effective if we understood the degree to which various diseases are rhythmic in occurrence. Similarly, accurate long range weather predictions, which would be of inestimable value in agriculture and aviation, depend on a knowledge of the nature and causes of the more or less rhythmic variations observed now and in the past.

Second, the techniques for dealing with rhythms and periodicities are much the same in one branch of science as in another, but unless the subject is considered a field by itself, advance in methods in one discipline is often unknown in another until after the lapse of many years.

Third, the subject of rhythmic and periodic fluctuations is important because identical rhythms in unrelated phenomena and unrelated fields of science suggest possible interrelationships that might otherwise escape notice. For instance, long before the cause of tides was understood, the fact that the period of the tides is exactly half the period of the moon suggested that the moon must have something to do with the ebb and flow of the waters.

These are some of the reasons why a concerted study of cycles as a whole is important. There are of course other reasons in respect to individual branches of science.

We must all be careful not to confuse contributions to our knowledge of *interrelationships* with contributions to the knowledge of rhythms and periodicities *per se*. There may be an excellent correlation between two phenomena, and yet no rhythms may be involved in either. The direct concern of The Foundation for the Study of Cycles is only with rhythms and periodicities as such. We must also keep in mind that *denial* of the existence of a rhythm may be as important as affirmation.

In addition to awarding a medal, the Foundation will make awards to outstanding work in each branch of science. In these awards the judges will be assisted by the recommendation of an advisor appointed by a scientific society in the branch involved. For the American Statistical Association Dr. W. Edwards Deming, of the United States Bureau of the Census, has been designated.

The person who receives the medal and the persons who receive the awards will also be elected Fellows of the Foundation for the Study of Cycles.

Communications should be addressed to Professor Ellsworth Huntington, Chairman of the Committee on Awards, Hendrie Hall, Yale University, New Haven, Connecticut.

A CORRECTION

In the memorial note of Norman John Silberling in the March 1943 issue of this JOURNAL, page 112, an error of omission was made. Dr. Silberling at the time of his death was Professor of Business Research at Stanford University.

BOOK REVIEWS

GLENN E. MCLAUGHLIN
Review Editor

Group Differences in Urban Fertility, by Clyde V. Kiser. Baltimore: The Williams and Wilkins Company. 1942. xii, 284 pp. $2.50.

This study, by a member of the Technical Staff of the Milbank Memorial Fund, forms an appropriate conclusion to an important series of investigations under private auspices directed toward the discovery and description of group differentials in fertility in the United States. Such investigations, on large scale and at high scientific level, were initiated by the same Fund a decade and a half ago. This study is the last of its kind, not because the problem can henceforth be neglected, but because its importance is now sufficiently demonstrated to be recognized as a matter of national interest and included in the agenda of the Federal Census. Private research in this field will henceforth be directed to related problems which are beyond the scope of census procedures, such as study of conditioning factors and motivation, and the elaboration of implications for social theory and policy.

It is apparent from Kiser's study and other investigations that the character of group differentials in fertility is undergoing an important modification. The general pattern of these differentials is not radically different today, but the monotonous iteration in the findings of various investigations is beginning to disappear. This is due, in part, to the development of more refined techniques; but it is also due in part to actual change in the phenomena themselves. As Ogburn has pointed out, differential fertility is in part due to lags in the diffusion of the family limitation pattern. So long as mere cultural distance from experimental science dominated the whole situation, the inverse relation of social status to fertility was fairly ubiquitous. In some cities, e.g., Boston, social gradients in fertility show little modification. This is true in general for all the large industrial cities in the Northeast. But in many parts of the country, group differentials in fertility show less uniformity and greater sensitivity to a variety of changing social conditions. Their accurate description will, therefore, require analysis of comparable data in different areas, under different conditions, and at different times.

Persons not concerned with the technical aspects of demography will generally await the publication of the forthcoming Census report on this subject for a description of the present status of group differences in fertility in the United States. This study by Kiser, however, explores some relations that will not be covered by the Census report. It also illuminates, in the course of the investigation, many other interrelations between occupational status, education, income, race, nativity, and locality.

Demographers will study this work primarily for its contributions to methodology. Many possible sources of bias in the interpretation of data on fertility and means of controlling such interpretations are fully developed

by the author and need no further elaboration. It seems to this reviewer that the limitations of marital fertility rates based on statistics of births in any limited period and the bias inherent in the standardization of marital fertility rates if interpreted as measuring total marital fertility are not sufficiently emphasized by the author. A standardized marital fertility rate necessarily eliminates the influence of differences in the span of the exposure period, which is one of the major factors influencing group differentials in fertility. Also, as the author points out, age-specific marital fertility rates, unless controlled for order of birth, are appreciably influenced by the length of the preceding interval since marriage, and the mean length of this interval at any given age varies widely among different groups. Similar bias, in lesser degree, affects standardized frequencies of number of children ever born to married women. Age-specific fertility rates supply valuable information on factors affecting fertility. However, the most accurate index of total marital fertility is obtained from data on children ever born to ever married women in age classes near the end of the childbearing period, and preferably with classification by criteria such as race, place of birth, education, etc., which remain constant or fairly constant through the whole childbearing period. This suggestion is emphasized here because of its possible bearing on census procedures; but it is likely to encounter some resistence from tidy statisticians who dislike the rejection of any available data that may seem on the surface to be relevant to the problem at hand.

The predilection of statisticians for the use of all collected data and the impossibility of carrying through all possible breakdowns with a limited body of material may be responsible for the use of the ambiguous category, "colored." The behavior, as regards fertility, of the Negro population is radically different from that of Mexicans, Chinese, and Japanese. There were many of the latter groups in the western cities included in the National Health Survey on which this study is based. Limitation of the data for nonwhite populations to those for Negroes, except where separate treatment of other ethnic-cultural groups was possible, would have obviated some of the author's constant difficulties with this material, somewhat analogous to those of a scientist dealing with zoo data on the feeding habits of seals, snakes, and "quadrupeds." This suggestion also has possible implications for census procedure.

The treatment of pregnancy wastage, which is in general a very useful contribution, is handicapped, as the author points out, by the gross underreporting of abortions and its possible influence on apparent group differentials. For example, the relatively high frequency of pregnancy wastage in the case of women with college education in the professional group (found to be statistically significant) might possibly be due to fewer inhibitions in this group against telling the truth on this subject.

An adequate description of differences in fertility in the United States requires a vast body of data specifically designed for this purpose. This is clearly shown by the difficulties which a skillful scientist encounters in the attempt to use for this purpose data incidental to another study. Dr. Kiser

and his predecessors and associates at the Milbank Memorial Fund are to be congratulated for blazing the trail, developing the necessary techniques, and demonstrating the possibility and significance of such studies. The Fund has already embarked on the still more difficult but no less important task of studying the conditions which determine such differentials and the possibilities of social control.

FRANK LORIMER

American University

Contraception and Fertility in the Southern Appalachians, by Gilbert Wheeler Beebe. Baltimore: The Williams and Wilkins Company. 1942. xii, 274 pp. $2.50.

When the millennium arrives with populations of optimum size maintained in every resource area of the world, the historians of that era should give due credit to the basic research in techniques and measurement of the control of human fertility which has been sponsored by the National Committee on Maternal Health and the Milbank Memorial Fund. Beebe's book follows closely the type of contribution made by Stix and Notestein in *Controlled Fertility*, with some further development in techniques for measurement and analysis of phenomena relating to contraception and fertility, and with adequate reporting of further experimental work.

Logan County, West Virginia, was the locale chosen for the three-year experiment reported. Within the county a group of women in the rural-nonfarm population, mostly wives of coal miners, were "enlisted" as contraceptive patients by a nurse working under supervision of local doctors. The technique advised was that of vaginal jelly alone with jelly and applicator supplied at no cost to the patient. For approximately 1,300 women who participated in the experiment, the chance of conception (as measured by number of second and subsequent conceptions per 100 woman-years of exposure to the risk of conception) was reduced by nearly one-half, from 73 to 40 for whites and from 82 to 45 for Negroes.

Such results form only the starting point of the detailed analysis presented. First is given an over-all review of the economic and social conditions of the Southern Appalachians of the sort that has been given before by regional sociologists and agricultural economists. Next follows a description of the project with particular attention to the problems of sampling involved. Like many other "experiments," this one began with scant concern for the representativeness of the individuals included in the study, but like few others, before the study was concluded, every effort was made later to explore their representativeness of rural-nonfarm women in the county. In fact, a major contribution to this line of research is Beebe's more careful consideration than has hitherto been evidenced of the importance of sampling procedures in the design of a study and of sampling errors in the drawing of conclusions from results.

The description of the group of women who participated in the experiment is interesting because they manifested in an exaggerated degree the traits common to the Southern areas of high fertility. Their mean age at first marriage was about seventeen and a half years, their admitted rate of premarital conceptions was 15 per cent for whites and 37 per cent for Negroes, their rate of first conceptions was such that 80 per cent had conceived within the first 10 months of marriage. From an auxiliary study which was made near the end of the experiment, it was found that about 32 per cent of the rural-nonfarm women of reproductive age in the county would accept contraceptive advice, but that these women because of higher fertility were responsible for about 50 per cent of the births of the county.

Hence the experiment demonstrates that similar programs on similar populations could reduce by about 40 per cent the fertility of a group of women who account for about 50 per cent of the county's births, or could reduce the total birth rate 20 to 25 per cent.

The conclusions and recommendations for the establishment of effective programs of contraceptive services again follow the Stix and Notestein lead of emphasizing the advisability of providing more than one technique to patients. This emphasis on the acceptability of a particular method to the patient rather than on the efficacy of a method under laboratory conditions has important implications for the planning of an actual clinic program.

Margaret Jarman Hagood

University of North Carolina

Public Health Statistics, by Marguerite F. Hall. New York: Paul B. Hoeber, Inc. 1942. xxii, 408 pp. $5.50.

In recent years there has been a virtual wave of new books on business statistics, social statistics, and general statistical methodology. Although the late Professor Pearl's third edition of *Medical Biometry and Statistics* appeared in 1940, the fact remains that new books devoted to vital statistics and biometry have been rare. This uncrowded condition of the field itself will probably ensure wide attention to Dr. Hall's *Public Health Statistics*.

The author states that this book "is intended for use as a text for a basic course in Public Health Statistics and as reference material in health departments." The introductory chapter contains a conventional but good discussion of history, principles, and limitations of statistical methodology. This is followed by three useful chapters on principles of computation and handling of mass data and tabular and graphical presentation of the materials. From this point the book can be said to fall into two major divisions, vital statistics and general statistics, although not so labeled. The vital statistics section includes Chapters V–XIV (pp. 88–259) respectively entitled population data; estimates of populations; standard millions; vital events; rates; mortality rates; vital statistics other than mortality; variation and trends in rates; ratios and index numbers; and construction of a life table. Chapters

XV–XXI (pp. 260–372), devoted mainly to general statistical methodology, are respectively labeled mass data—preliminary definitions; measures of central tendency; measures of variability; measures of relationship and trend; theory of probability, point binomial, normal curve; measures of reliability, and prerequisites to the application of statistical methods.

The author probably had good reasons for the tandem arrangement of the materials for vital statistics and those for general statistics, but the rewiewer fails to see what they were. If separation was necessary it would seem that the chapters on vital statistics should follow rather than precede those on general statistics. In the reviewer's opinion, however, better integration of the materials would be in order for a book on public health statistics. As one might expect, the present arrangement resulted in virtual duplication of the discussion in several instances.

The author avowed a conscious effort to prepare a book that could be used by people with little training in mathematics. With commendable simplicity of language and detail she leads the reader through the step-by-step procedures in specific computations. Dr. Halbert L. Dunn correctly states in the foreword, "She has discovered the knack of making probabilities, chi-square rates and index numbers into everyday county, city, and state health language." This characteristic will doubtless enhance the value of the volume as a handbook for use in health departments.

In the reviewer's opinion, however, the volume does not measure up so well as a textbook. For this purpose one would look for adequate discussion of underlying theory as well as for clarity in the exposition of technique. Unfortunately, the discussion of theory tends to be sketchy. It is possible that in writing the book the needs of the health department rather than those of the classroom were uppermost in the author's mind. Whatever the explanation, the author tended to devote much more attention to the mechanics of computation than to the meaning of the statistical measures considered. On several occasions the reviewer was somehow reminded of the lines, "Theirs not to reason why, Theirs but to do and die."

Despite the above characteristics there is ample evidence of the author's own well-balanced point of view regarding the place of statistical technique in public health work. She emphasizes strongly the limitations of statistical methodology and the necessity of knowing at first hand the characteristics of the basic data and the conditions under which they were collected.

CLYDE V. KISER

Milbank Memorial Fund

Refugee Settlement in the Dominican Republic, A Survey Conducted under the Auspices of the Brookings Institution. Washington: The Brookings Institution. 1942. xvi, 410 pp. $4.00.

In 1938 at an intergovernmental conference on the refugee problem, the Dominican Republic, alone among thirty-two nations, declared its willingness

to welcome large numbers of refugee settlers. The limit was set at 100,000. This offer, the one bright spot in a gloomy picture, speedily led to the formation of the Dorsa, the Dominican Republic Settlement Association, a social agency financed from private funds, and to the establishment on the north coast of the Dominican Republic of the first refugee colony of under 500 persons at Sosua, a 26,000-acre estate offered to the Association by Generalissimo, now again President Trujillo.

This survey is limited to the activities and prospects of the Dorsa; it does not extend to other refugees, among whom Spanish loyalists are particularly numerous. Moreover, the survey confines itself strictly to economic aspects of the problem, omitting social, political, and other non-economic phases. Nevertheless, this book is far more than the study of a single local settlement. It is the most comprehensive study of the economic life of the people of the Dominican Republic known to the reviewer and as such constitutes a worthy companion volume to the Brookings Institution survey *Porto Rico and Its Problems*, 1930. The success or failure of the Dorsa and its protegés is inextricably tied up with the economic life of the Republic.

This economic survey of the Dominican Republic makes up the bulk of the study. It is preceded by three chapters dealing with the refugee problem in general and is followed by four chapters devoted to the work and problems of the Dorsa in particular.

The principal questions which the survey had to consider were: (1) What is the relation of the Dominican project to the refugee problem as a whole? (2) Is the Dominican Republic a suitable place for refugee settlement? (3) How may refugee immigrants make a living there? (4) How many settlers can the Republic absorb? (5) What will be the effect of the settlement project on the Dominican community?

Of these five questions the third and fourth are of most immediate and practical importance. The answers given are none too encouraging. Refugee immigrants may aspire to the level of peasants, i.e., agriculturists devoting most of their energies to subsistence farming with cash crop farming or primitive industry definitely supplementary and secondary. The Republic can absorb according to the majority of the survey staff from 3,500 to 5,000 persons and, according to one member, up to 10,000 refugee settlers.

Those familiar with the desperate conditions prevailing throughout the Caribbean area, the answer to question three can hardly be surprising. But the low absorption capacity of the Dominican Republic does come as a shock to those who compare the Dominican rural population density of 68 per square mile with that of Puerto Rico, Haiti, and Jamaica, which in the aggregate is over four times as high. The low arability coefficient, the fantastic rate of increase of the population, and the land requirements of the refugee peasants, which far exceed those of the native rural proletariat, offer the explanation.

It will cost not much less than $4,000 to settle a refugee on the island and of that he is expected to pay back about $1,600. At that rate the settlement

of 5,000 "units" would require an outside net contribution of roughly 12 million dollars.

Apart from the chapter on the history of the Republic which reads like a dime novel, the book is written in a sober, almost sombre tone. It is an invaluable study of a local problem with global implications. It is required reading for all those wrestling with the problems of the post-war world.

The map of the Island of Hispaniola facing the title page might have been enhanced in usefulness by the insertion of names prominently mentioned in the text. The absence of an index is not a serious flaw in a survey carefully outlined and skillfully articulated.

ERICH W. ZIMMERMANN

The University of Texas

Consumer Instalment Credit and Economic Fluctuations, by Gottfried Haberler. New York: National Bureau of Economic Research. 1942. xix, 239 pp. $2.50.

This volume is the ninth of the studies in consumer instalment financing prepared under the financial research program of the National Bureau of Economic Research. In it data gathered in the previous studies are used as the basis for a theoretical analysis of the relations between consumer instalment credit and cyclical fluctuations in general business.

Consumer instalment credit both in the form of commodity credit and cash loans shows a pronounced cyclical pattern when measured by any one of the series used for this purpose, volume of new credit granted, amount of credit outstanding and "net credit change." The causes of fluctuations in instalment credit are found in both long run growth factors as well as cyclical factors. Cyclical fluctuations in consumer credit are caused more by cyclical shifts in demand than by cyclical shifts in supply. Responsiveness of demand to changes in the terms of credit is more dependent upon changes in the down payment percentage and contract length than upon changes in the rate of interest charged.

The introduction and expansion of consumer instalment credit in the 1920's had a stimulating effect upon the economy in the short run. Whether this development is to be regarded as favorable or unfavorable for economic stability in the long run depends upon questions of business cycle theory on which expert opinion is still divided. Haberler's conclusions are as follows. After the introduction and development to maturity of instalment credit as an institution it is still subject to cyclical fluctuations but these fluctuations are then the result rather than the causes of cycles in general business activity. However, consumer credit does act as an amplifier by intensifying the movements of the general cycle whether up or down. Yet this behavior is in no way different from that of credit in general.

In comparison with other influences upon general business activity consumer instalment credit was not very important during the decade 1929–40,

much less important, for example, than the financial operations of the Federal Government. Yet consumer credit has a greater significance than this comparison would suggest when it is considered in relation to the development of particular industries and their effects upon the economy as a whole. Consumer credit as an institution should be maintained, but efforts should be made either to mitigate its cyclical fluctuations or to bring about an anti-cyclical pattern.

The publication of this book is timely because of the new regulations of consumer credit of the Federal Reserve System in the present emergency and the possibility that such regulations may be carried over into the post-war period.

James F. Cusick

Dartmouth College

Farm-Mortgage Credit Facilities in the United States, by Donald C. Horton, Harald C. Larsen and Norman J. Wall. Washington, D. C.: United States Government Printing Office. Miscellaneous Publications No. 478. United States Department of Agriculture. 1942. vii, 262 pp. 55 cents.

A more appropriate name for this publication would have been Farm-Mortgage Credit Facilities *of* the United States, Why they should be increased and How they can be increased. Nine of the sixteen chapters deal exclusively with federal credit agencies and the remainder consider the more general problems of debts, interest rates, etc., with but little mention, let alone description or analysis, of other farm-mortgage credit facilities. Yet, the following statement expresses a point of view which permeates the entire discussion.

> But the public purposes to be achieved through extensive participation of the Government in farm-mortgage financing may be so essential as to make the disadvantages negligible. It is possible that the results expected of a co-ordinated Federal mortgage-credit service are such that the profit motive of private financing alone cannot be expected to accomplish them. For example, public policy operating through Federal mortgage-credit channels might be made to serve as a beneficial form of public action designed, along with other action programs, to assist agriculture in adapting to changing economic conditions. Considerations of this character may justify the sacrifice of certain of the advantages to farmers of private credit facilities. (Page 170.)

The above can hardly be called a conclusion because it is not based on any evidence in the text. It is not stated as an assumption on the basis of which further argument is advanced. It does, however, express a philosophy which is evidenced in numerous parts of this work. The reviewer has no wish to quarrel with the authors for this point of view. He does wish to point out, however, that evidence is not presented to support it other than evidence which indicates that the Federal Government did expand its credit facilities

in the depths of depression. The fear expressed is that unless these agencies adapt themselves to present economic conditions they may not be available to serve should similar catastrophies arise in the future. Then, of course, the happy thought occurs that such a calamity may be avoided if "wise" credit policies are followed.

The authors maintain that "if the Federal agencies are to be major sources of farm-mortgage credit in periods of prosperity as well as in depression, it may be necessary to modify these institutions further so they can serve not only as refinancing and financial-adjustment agencies but also as continuous sources of both purchase-money and operating credit."

Though the authors may seem to the reader to be advocates of "socialized credit" they do advocate the sound policy of modifying by improvements existing Federal farm-mortgage credit facilities to meet the varied needs of agriculture under varied conditions. As a basis on which such modifications may be made there is presented a wealth of information, tabulated in a very orderly manner and unusually well illustrated.

Part I discusses the historical background of the farm-mortgage credit situation for the period 1910 to 1940. Though the general theme is much the same heard so many times, particularly during the last decade, its concise recording with suitable statistics is a useful adjunct to the text.

Part II is particularly good in its description of Federally sponsored farm-mortgage loan agencies. The organization and development of each agency is described and its source of funds, terms and conditions of loans explained.

The first two parts carry out very well the intentions of the authors, namely "to serve a principal purpose of providing a compact summary and analysis of factual material on farm-mortgage credit heretofore available only in scattered sources."

Part III is concerned with current problems and issues in farm-mortgage credit policy but as indicated in the introduction "more emphasis is placed upon identification of problems and issues, and upon exploration of their interrelationships in the light of alternative public objectives than upon formulation of specific recommendations."

The publication merits careful study by all borrowers and prospective borrowers and by lenders, both public and private, at this particular time when foundations are being built for the economic structure within which we must all live for the next decade or two.

JOHN B. SCHNEIDER

University of California, Berkeley

Business Behavior 1919–1922, by Wilson F. Payne. Chicago: The University of Chicago Press. 1942. xi, 215 pp. $1.50.

This book must be regarded as timely, since many people are now interested in accounts of past experiences of the postwar development of business which may be used for the purpose of forecasting the future development of business after the present war. The approach of the author to his

subject is more than empirical; it may be qualified as "atheoretical." In the introductory chapter, he complains of too many contradictory theories of business fluctuations and points to the danger of selecting some one hypothesis as a basis for analysis of data, since this selection of itself determines what kind of data is considered as important in such an analysis.

The author follows an industry approach to his subject. Several industries were selected for survey, most of which are divided into a number of integral segments. He analyzes leather and shoes, textiles, iron and steel and their principal users, automobiles, and building and allied industries. The last group is not included for close examination, but simply to complete the general picture of the period. These surveys are preceded by a short analysis of the financial situation during the period under study. The author recognizes that this aspect of business development during 1919–22 has received keen examination by other observers, but he found it advisable to review it in order to provide necessary points of reference.

The author thoroughly worked numerous statistical data, which are summarized graphically in eleven charts, each including some 15 to 20 curves. These charts represent the behavior of the various indexes characteristic of the industries surveyed in the book. But they are not interpreted in the text. This may be regarded as a shortcoming from the point of view of the general reader, since comparisons and interpretations of these charts are not easy problems. Only a specialist may do so advantageously. In the opinion of the reviewer, this omission, which presumably was done for economy of space, may be partly responsible for the fact that Payne's book does not leave the reader with an impression of the behavior of business as a whole. Although the interactions between segments of individual industries are sufficiently analyzed in separate chapters, little attention is given to a systematic analysis of the interactions between the major branches of industry during the 1919–22 cycle.

The last chapter gives a recapitulation of developments, but it occupies only nine pages, and its conclusions do not utilize fully the rich material supplied in the separate chapters. The lack of sufficient analysis of interconnections may result from the atheoretical approach of the author. He purposely avoids using any of the analytical tools elaborated by the theory of the business cycle, as if he is afraid of falling in the tenets of some one theory. But his statement that the list of industries selected for analysis "is confined to a representation of those of which short industrial depressions are peculiarly composed" presupposes some theoretical conception of the business cycle, which he does not formulate explicitly. By excluding the food industries and agriculture, he disposes of at least one group of the theories of the business cycle. The exclusion of agriculture is not consistent, however, since in his analysis of the textile industries he was obliged to pay much attention to variations of cotton crops and his analysis shows how important the quantitative and qualitative characteristics of these crops during the years under study were for the explanation of the business situation in the cotton industry.

The extreme empiricism of the author, however, may be justified by the unsatisfactory state of the theory of the business cycle. In its present situation, the theory may be regarded more of a barrier than a help to a theoretico-empirical approach of the subject, and may contribute to extreme empiricism of even those investigators who are conscious of the dangers of extreme empiricism, as is the author.

The book may be used advantageously in advanced courses on the business cycle, particularly in special seminars, though the general reader will hardly get a general picture of the business behavior during 1919–22. Those interested in the behaviors of separate industries during that period will find the book interesting and useful reading.

V. P. TIMOSHENKO

Food Research Institute
Stanford University

Commodity Statistics, The 1942 Commodity Year Book. New York: Commodity Research Bureau. 1942. 294 pp. $5.00, and *Commodity Year Book.* New York: Commodity Year Book. New York: Commodity Research Bureau, Inc. 1942. 413 pp. $7.50.

The volume on *Commodity Statistics* is the fourth of an annual statistical series. Descriptive materials on the background, processing, and marketing of commodities have been omitted from the 1942 volume. Such information covering all the commodities listed in the statistical volume and in addition virtually all of the important crude and semi-finished products used in the national economy are presented in the *Master Edition, Commodity Year Book.* The latter, therefore, is a commodity encyclopaedia designed for permanent reference use.

Commodity Statistics presents data in 61 individual or group sections for the 75 important crude and semi-finished materials covered in previous statistical editions. Statistics reported are for commodities generally classified as animal and animal products, beverages, forest products, fuels and electric power, metals, nonmetallic minerals, vegetable foodstuffs, vegetable oils and fibers, and such specialties as rubber, silk, and tobacco.

Data presented in each section vary in length from 1 or 2 pages to 16 pages depending on available information and the importance of the commodities. The sections generally comprise statistics on consumption or sales, production, prices, exports and imports, and stocks. Frequently statistics are given on such subjects as yields, taxes, kinds of uses, and sizes. All the statistical tables are simply arranged and easy to read. Footnotes are given but in some tables, particularly those involving estimates and comparability of a series of data, technicians may desire more detailed explanations.

The data have been compiled from a variety of recognized sources such as Government agencies, Boards of Trade, commodity exchange bureaus, and other institutions. Source references to the pertinent agencies and institu-

tions are printed at the foot of each table. This source information could be more helpful to statistical workers if it also included references to the names of the current publications of the particular agencies. Technicians using the volume may find it necessary to look up available current data.

Information in the 1942 Year Book has been somewhat clouded by war conditions. Only data that have been officially released by governmental and quasi-official sources as not being detrimental to the war effort have been published. Consequently, statistics for strategic and critical commodities generally are not shown after the first 9 to 11 months of 1941.

The *Master Edition* is composed of separate articles on over 800 different crude and semifinished commodities. These articles contain brief descriptions of available information on such subjects as physical characteristics, methods of production, marketing and transportation methods, comparative prices over recent years, principal types and grades, possible substitutes, and government regulations. An attempt was made to use the same general outline for all commodity articles, but variations are frequent because such data for many commodities are lacking in part or otherwise do not warrant such specific treatment. Amount of material presented for each commodity varies from part of one page for the less used commodities up to 16 pages for a group of commodities such as plastics. Notable features of this volume are its brief historical and economic reports on rubber, plastics, and many common and rare chemical and vegetable commodities.

The two volumes should be valuable to statisticians, investigators, and research workers, because extensive statistical and informational data for so many important commodities are to be found in one set of reference books. It is hoped that the editors will find it possible and necessary to expand the coverage of their statistical volume to include some data for all commodities described in the *Master Edition*.

Oscar L. Endler

National Resources Planning Board

Make or Buy, by James W. Culliton. Boston: Harvard University, Bureau of Business Research. 1942. iv, 130 pp. $1.50.

Every business purchase, the author points out, involves at least a tacit decision that it is more desirable to buy the article in question from an outside supplier than to produce it within the organization. Conversely, every production schedule indicates at least a tacit decision that each item involved can more satisfactorily be made by the company than purchased outside.

While there are limits to the field of such choices, the author believes that too often a decision to make or buy a particular item, regardless of its original merit, continues to stand as an established policy long after changing conditions call for further review. Obviously, few companies will find it economical to make their own lead pencils or stationery; a specialty manufacturer

will not seek to put himself out of business by buying the finished articles which provide his primary manufacturing revenue. Between such limits, however, lies a wide range of optional items which could be either made or bought.

Cost will be the principal factor determining such a decision, the author emphasizes. He advocates no one method of determining costs, pointing out that the facts will vary from case to case in the matter of overhead involved in producing a particular part, in the length of tool life, and in optional uses of existing facilities. Savings or increases in purchasing expense and storage and handling costs may or may not be capable of determination and allocation for individual items. In the analysis of a number of case studies, the author points out that all costs that can be clearly allocated should be, and that the decisions reached in some cases might have been reversed had a different analysis of cost been made. The final decision, even where cost is concerned, depends to a large extent upon the exercise of business judgment.

Further factors which will influence "make or buy" decisions, particularly where the cost element is not decisive, are control over quality, control over scheduling of production, and risks beyond the company's control. The author questions whether these factors deserve as much weight in decisions not to buy as they often receive, and cites cases to support his position.

Readers in search of clear-cut generalizations as to the preferability of making or buying will be disappointed. The analysis and the wealth of case material from a variety of manufacturing operations will, however, make the book of interest to individuals reviewing their own company policies. With many companies facing postwar reconversion and a fresh start in their regular peacetime operations, the book is particularly timely.

CHARLES E. YOUNG

Westinghouse Electric and Manufacturing Company

Wool and the War, by Hiram S. Davis. Philadelphia: University of Pennsylvania, Wharton School of Finance and Commerce, Research Report No. 6. 1942. xi, 107 pp. 50 cents.

This publication presents a clear picture of our war supplies problems for wool textiles. The author carefully weighs the wool supply prospects compared with probable needs for the war period. The report analyzes stock on hand and prospective production and importation of raw wool, prospective production of substitute materials, and the country's capacity for producing wool textiles as measured by mill facilities and trained labor. Wool needs are appraised in terms of prospective military and civilian requirements through 1944.

The author concludes that, although our present and prospective supply of raw wool may not be enough to support civilian consumption at its peacetime rate without encroaching upon military reserves, there is no foreseeable shortage of substitute fibers ahead. Continuance of the use of raw wool for

civilian products at its current restricted level (the conservation order of the WPB) has limited civilian use to about 44 per cent of annual peace-time consumption, and the difference between this volume and usual peace-time consumption the author contends can easily be made up by using in apparel production more of the wastes from wool processing, more wool fiber recovered from tailor's clips and old rags, more hair fibers, and finally more blends of wool and rayon or cotton, or of cotton-warp wool goods.

The author warns that this plentiful supply of substitutes should not mean that consumers can expect to buy wool textile products at the high level of 1942 without sooner or later creating serious shortages. This warning note seems very appropriate in view of the prospective increase in purchasing power in 1943 and 1944, coupled with the expanded rationing program of non-clothing items which concentrates more consumer buying pressure on clothing items.

The report concludes that the maximum military and civilian demand for wool textile products, which now seems in prospect, could be met from existing mill facilities if they were operated on a three-shift rather than on a two-shift basis. Even with two shifts of 40 hours per week, however, the report concludes that the industry could handle a military business equivalent to the total of 1942 and a civilian output of apparel, blankets, and woven felts very near the level of 1939.

According to the author both materials and productive facilities should be adequate for prospective military and civilian demand. It should be remembered, however, that this report was written during the first half of 1942 and events which have occurred since that time have changed significantly some of the prospects particularly on the demand side. For example, the author bases his estimates of American supply on a maximum of five million men in the armed forces at the end of 1942, seven million by the end of 1943, and nine million by the end of 1944. Most recent estimates indicate at least eleven and a half million will be in the armed forces by the beginning of 1944, or shortly thereafter. This increase in demand, coupled with continued heavy buying of wool textiles, makes the wool supply picture less bright. Partly offsetting this greater demand situation is the record 1942 wool production of 459,073,000 pounds, which is an increase of nearly sixty-nine million pounds over the 1941 output. However, the author concludes that "the sheep expansion program does not offer significant results either from the point of view of the immediate emergency or from that of nationalism." In support of this conclusion, the author cites the failure of higher wool prices during World War I to result in any appreciable increase in sheep numbers. Despite the tripling of prices, the domestic clip in 1919 was only five per cent above that of 1914, and the wool clip would have shown an actual decline had it not been for expansion in Texas production. It is significant that sheep numbers in the United States at the beginning of 1943 are about three per cent below those a year earlier.

The report is well buttressed with ample supporting factual information

and a thorough knowledge of the wool textile industry. Moreover, the author is aware of developing events because he ends the report with a warning note that rationing wool textiles may be necessary not because wool textile supplies may fall short of ordinary needs, but because unspent purchasing power may push demand up to extraordinary levels. Also he points out appropriately that neither the discouragement nor rationing of consumer purchases really solves the problem of existing purchasing power, and that some steps must be taken to remove at least a portion of existing consumer purchasing power from the markets. Heavier taxation, enforced savings, and related fiscal controls are essential if purchasing power is to be kept down to a point where consumer sales will not create severe shortages of most wool textile materials. The implications of these requirements on national fiscal policy and the OPA rationing program should be obvious.

R. R. RENNE

Montana State College and
Office of Price Administration

Wartime Rationing and Consumption. New York: Columbia University Press. League of Nations Publication. 1942. II. A. 2. 87 pp. $1.00.

As its preface states, this short book "is concerned mainly with the impact of war controls and rationing on consumption and standards of living in countries for which information is available."

Its first brief chapter treats of methods and principles of consumption control and of direct consumer rationing. An important observation is that wartime governmental rationing systems " . . . contain the elements of a distributive system in which consumption is guided not so much by individual purchasing power as by human wants." It will be interesting to see whether postwar developments nurse or strangle this growing plant.

The long second chapter presents statistical evidence of food rations in many countries as of April 1942. For understandable reasons the sources of data cannot be given, and therefore the results cannot be checked by reviewers. The wartime history of changes in rations is touched briefly, and comparisons are drawn between the 1942 rations and prewar consumption. The value of this compilation and comparison, which represents immense labor, goes almost without saying; it constitutes the most comprehensive survey presently available. Its value is not vitiated by the following comments: Rumanian (Bucharest) rations of sugar, meat, and fats look incredibly high. The conclusion that Italy ranks with France, Belgium, Norway, Finland, and Spain as having rations so abnormally low as to imply prevalence of both undernourishment and malnutrition is difficult for the reviewer to reconcile with such advices as seep out of the country. The section on "Worker's Family Consumption" carries little conviction so far as concerns comparisons of rations with prewar calorie consumption as shown by budget surveys, because the range of calorie consumption indicated by the surveys

was incredibly large. Acceptance of these surveys, and of various nutritional standards advanced by the League of Nations, seems somewhat uncritical.

The very brief third chapter gives a worldwide survey of wartime rationing of clothing, housing and household essentials, miscellaneous necessities and services, and non-essentials. The final chapter seeks to measure the effects of wartime spending restrictions upon the aggregate volume of consumption in seven countries, especially the United Kingdom, Germany, and the United States.

Readers interested in finding a guide for American policy may question the emphasis upon the desirability of equity of food distribution in a full-fledged war economy. They may also feel a lack of discussion of the labor costs of consumer rationing, of the impact of rationing on the distribution system, of the circumstances of general or only of particularized food shortages that justify food rationing, and of the difficulties of holding farmers within food-ration allotments.

The subject matter—principles and practices alike—lends itself to continuing inquiry and development. It is to be hoped that supplementary publications will appear.

M. K. Bennett

Food Research Institute
Stanford University

Statistical Methods Applied to Agricultural Economics, by Frank A. Pearson and Kenneth R. Bennett. New York: John Wiley & Sons, Inc. 1942. vii, 443 pp. $4.00.

This is a valuable book for the practical agricultural economist. It combines in a single volume nearly everything of a practical nature which the average student of agricultural economics needs to know about statistical method. The authors have illustrated the various statistical techniques with numerous examples taken from practical problems largely in the fields of farm management, marketing, and prices. Statistical theory is subordinated to its practical application to real problems, but the student should easily get the theory from study of the specific examples to which it is applied. The language of the text is fairly non-mathematical. The interest and attention of the agricultural economist reader will be held by the wide variety of well chosen practical problems which illustrate the statistical procedures discussed.

The general arrangement of the book follows the usual order of statistics textbooks: measures of central tendency, variation, association, and reliability. It is divided into 22 chapters. Nearly every chapter has a very helpful section near its end entitled "Uses," and/or "Comparison," in which the authors summarize in a non-mathematical and practical manner the meat of the chapter.

Two excellently prepared chapters on the tabular analysis of relationships cover a field which has been neglected entirely too long in agricultural statis-

tics books. This method has been and will continue to be the most widely used method of analyzing relationships. Perhaps its general usage and apparent simplicity have prompted statistical writers to dismiss it lightly. Yet it is well for the student of agricultural economics to know its uses and limitations. The manner in which the authors develop the treatment of tabular analysis provides an easily understood introduction to the theory of correlation in the following chapter.

Regression and correlation are treated in a more understandable manner than is true in many textbooks. There is considerable material of a practical nature in these chapters that is frequently omitted from books of this kind.

Some readers may feel that the authors have placed too much emphasis upon the analysis of data already collected, and too little emphasis upon the theory of sampling and the planning and collection of an adequate and representative sample. Agricultural economists too frequently follow some variation of the "hit or miss" method in the collection of sample research data, and then go to great pains to be "accurate" in its summarization. It is essential that agricultural economics students and researchers alike be made increasingly aware of the importance of proper sampling technique.

Other readers may notice the omission of an introductory chapter explaining the general relationship of statistics to science and research. Statistics are a very useful tool, and it may seem that any statistics textbook should attempt to orient the beginning student in an appreciation of the contribution that statistics can and should make to the advancement of knowledge.

E. L. BUTZ

Purdue University

Fundamental Statistics in Psychology and Education, by J. P. Guilford. New York: McGraw-Hill Book Company, Inc. 1942. xi, 333 pp. $3.25.

The intention of the author in preparing this book is to present the basic principles and methods of statistics to the beginner in the subject. The treatment is essentially non-mathematical in character, although mathematical symbolism is freely used.

No attention is given to methods of collection and arrangement of data. The major emphasis is upon methods of analysis and the interpretation of the results of analysis. Early chapters deal adequately with measures of frequency distributions, and the properties of the normal curve. The remainder of the book is devoted to the problems of sampling, measures of reliability of samples, and correlation procedures.

Perhaps the outstanding feature of the book is the uniform clarity of presentation. There are certain points at which the reviewer feels that the fetish of brevity may have been pursued at the expense of clarity and completeness, but such lapses are negligible in comparison with the overall precision of expression and adequacy of treatment.

This matter of brevity is most prominent in the last two chapters of the

book. The chapter on multiple and partial correlation would be more satisfying if an explanation were included of the meaning of multiple and partial regression equations. Some explanation is also needed of why the successive steps are taken in the Doolittle Method of solution. The chapter on reliability and validity of tests is an excellent critical discussion of the theoretical problems involved in evaluating test results. It seems to imply more practical experience with tests than the beginning statistical student would ordinarily be presumed to possess.

At the end of each chapter a set of exercises is appended. These exercises are well designed to reinforce the student's understanding of the principles set forth in the chapters, and are adequate for homework assignment or for a laboratory of moderate length.

The attitude of students toward courses in statistics has in many cases become extremely adverse because the authors of textbooks have succeeded only in introducing mystery into the subject. This book takes the mystery out of the subject by its clear and rounded exposition, its judicious choice of illustrative examples, and its general adherence to the point of view of the beginning student. The book is indeed a valuable teaching medium.

Martin A. Brumbaugh

University of Buffalo

An Outline of Probability and Its Uses, by Maurice C. Holmes. Minneapolis: Burgess Publishing Company. 1936. viii, 119 pp. $1.50.

Four decades of experience in a great industrial organization justify this reviewer in stating that the *Outline* by Professor Holmes should be in the hands of every engineer and statistician unfamiliar with the theory of probability. Moreover, the *Outline* offers an excellent résumé for one already acquainted with the subject

Commendation is called for by the author's choice of material and the order of its presentation. The diagrams throughout the text will appeal particularly to the engineer.

In this time of stress, when our thoughts should be concentrated on winning the war, the above should suffice as a review. However, the following comments are added in anticipation of a second and other editions, and as proof that the reviewer has read the first edition from a to z.

1. On reading page 7, where mutually exclusive events are dealt with, the novice may interpret "the probability that *at least* one will happen" as meaning that "not less than one will happen." It might be clearer to say "some one of the events" or "some one or other of the events will happen."

2. On page 10, no doubt the author is historically correct in his statement that equation (9) is "called the binomial law from the fact that the term C_n^m is equal to the coefficient of the nth power of x in the binomial expansion of $(x+y)^m$." In our industry we apply the name "binomial law" to equation (9) because its entire right hand term is a term in the expansion of the binomial $(p+q)^m$, wherein $q=1-p$.

3. On page 99, under the caption "Symbols," the equation

$$a+b = \text{"}a \text{ or } b\text{"}$$

might better be written

$$a+b = \text{"}a \text{ or } b \text{ or both."}$$

One unfamiliar with symbolic logic may interpret "a or b" as meaning "*a only* or *b only*."

The above comments in no way reflect on the merits of the *Outline*. The exposition goes straight to the point without any lost motion. Moreover, it is entirely free from facetious phrases, a type of phrase which disfigures many otherwise excellent works.

E. C. MOLINA

Seven-Place Values of Trigonometric Functions, compiled by J. Peters. New York: D. Van Nostrand Company, Inc. 1942. $7.50.

Those acquainted with the well known tables *Siebenstellige Werte der trigonometrischen Funktionen* compiled by Dr. J. Peters of the Kgl. Astronomische Rechen-Institut, Berlin, and published in 1918, will welcome this photo-offset of those useful tables together with the English translation of the Introduction. The Alien Property Custodian has undoubtedly performed a valuable service in making these tables generally available at this time.

The need for increasing the speed of computations and the increased use of electric calculating machines in astronomical, geodetic, and optical calculations has emphasized the value of accurate natural trigonometric functions. These tables admirably fill that need. The use of decimal parts of degrees is logical in simplified calculations. Since these tables are photo-offsets they will not be subject to typographical errors as new type-set tables would be. This reviewer has used the original tables with the electric calculator for over a decade finding them more satisfactory than any others currently available.

These tables which give the sine and tangent (cosine and cotangent) in thousandths of a degree and the value of the function to seven places are based upon those of Gellibrand's *Trigonometria Britannica*, (Goudae 1633) which were computed for every hundredth of a degree of the quadrant. The functions in the present tables were interpolated for the thousandth of a degree, except for the tangents of angles from 89° to 90° where the values were calculated. The possible maximum error is given as ± 0.65 unit in the seventh decimal place of the sine function and as ± 0.85 unit in the seventh decimal of the tangent function. In general the errors are very much less than these and interpolations may be made for a ten-thousandth of a degree with considerable accuracy. Conversion tables of minutes and seconds to decimal parts of degrees and of degrees to time (and vice versa) are given at the end of the book. The tables are arranged for convenient use.

KENNETH N. OGLE

Dartmouth Eye Institute
Hanover, New Hampshire

PUBLICATIONS RECEIVED*

Barger, Harold, and Landsberg, Hans H. *American Agriculture, 1899–1939: A Study of Output, Employment and Productivity.* New York: National Bureau of Economic Research. 1942. xxii, 440 pp. $3.00.

Bennett, M. K., Farnsworth, Helen C., and Peirce, Rosamond H. *Wheat in the Third War Year: Major Developments, 1941–42.* Stanford University, California: Wheat Studies of the Food Research Institute. Vol. XIX, No. 3. 1942. 35 pp. $1.00.

Bernays, Edward L. *The Future of Private Enterprise in the Post-War World.* New York: Counsel on Public Relations. 1942. 11 pp.

Bowley, A. L. *Studies in National Income.* Cambridge: The University Press, New York: The Macmillan Company. 1943. ix, 255 pp. $3.50.

Bowman, Mary Jean, and Bach, George Leland. *Economic Analysis and Public Policy.* New York: Prentice-Hall, Inc. 1943. xix, 935 pp. $4.50.

Clark, Fred E. *Principles of Marketing.* New York: The Macmillan Company. 1942. xx, 828 pp. $4.50.

Culliton, James W. *Make or Buy.* Boston: Harvard University, Bureau of Business Research. 1942. iv, 130 pp. $1.50.

Davis, Joseph S. *New International Wheat Agreements.* Stanford University, California: Wheat Studies of the Food Research Institute. Vol. XIX, No. 2. 1942. 58 pp. $1.00.

Fabricant, Solomon. *Employment in Manufacturing, 1899–1939.* New York: National Bureau of Economic Research. 1942. xix, 362 pp. $3.00.

Fisher, R. A. *The Design of Experiments.* Edinburgh: Oliver and Boyd Ltd. 1942. Third Edition. xi, 236 pp. 12 shillings 6 pence.

Frickey, Edwin. *Economic Fluctuations in the United States.* Cambridge: Harvard University Press. 1942. xxi, 375 pp. $5.00.

Gras, N. S. B. *Harvard Co-operative Society, Past and Present.* Cambridge: Harvard University Press. 1942. xi, 191 pp. $1.50.

Hardy, Charles O. *Do We Want a Federal Sales Tax?* Washington: The Brookings Institution. 1943. 47 pp. 25 cents.

Hoyer, Louis P., and Hay, Charles K. *Services to the Orthopedically Handicapped.* Philadelphia: Walther Printing House. 1942. xv, 115 pp.

Hurlin, Ralph G. *Salaries and Qualifications of Child Welfare Workers in 1941.* New York: Russell Sage Foundation. 1943. 27 pp. 25 cents.

International Labour Office. *Year Book of Labour Statistics.* Montreal: International Labour Office. 1943. 222 pp. $3.00.

Kurtz, Russell H., Ed. *Social Work Year Book, 1943.* New York: Russell Sage Foundation. 1943. 764 pp. $3.25.

League of Nations. *Economic Fluctuations in the United States and the United Kingdom, 1918–1922.* New York: Columbia University Press. 93 pp. $1.50.

Lombardi, John. *Labor's Voice in the Cabinet.* New York: Columbia University Press. 1942. 370 pp. $4.00.

Magill, Roswell. *The Impact of Federal Taxes.* New York: Columbia University Press. 1943. ix, 218 pp. $3.00.

Mendershausen, Horst. *The Economics of War.* New York: Prentice-Hall, Inc. 1943. xiv, 390 pp. $3.35.

Merwin, Charles L. *Financing Small Corporations in Five Manufacturing Industries, 1926–36.* New York: National Bureau of Economic Research.

* Annual reports and publications presenting statistics collected at regular intervals have been omitted from this list. Some items of minor interest to statisticians have also been omitted. The contents of periodical publications are not listed, but the attention of the reader is directed to the lists of articles in current publications which are to be found in the *Revue de l'Institut Internationale de Statistique, Journal of the Royal Statistical Society, American Economic Review, Population Index, Transactions of the Actuarial Society of America, The Record of the American Institute of Actuaries,* and *Sankhyā, The Indian Journal of Statistics.*—Editor.

Studies in Business Financing. 1942. xvi, 172 pp. $1.50.

Metz, Harold W. *Is There Enough Manpower?* Washington: The Brookings Institution. 1942. 25 pp. 25 cents.

Mitchell, Wesley C. *Wartime 'Prosperity' and the Future.* New York: National Bureau of Economic Research. Our Economy in War, Occasional Paper 9. 1943. 40 pp. 35 cents.

Moulton, Harold G., and Schlotterbeck, Karl. *Collapse or Boom at the End of the War.* Washington: Brookings Institution. Pamphlet No. 47. 1942. 40 pp. 25 cents.

National Bureau of Economic Research. *Income Size Distributions in the United States, Part I.* New York: National Bureau of Economic Research. 1943. xxvi, 131 pp. $1.00.

Nourse, Edwin G. *Democracy as a Principle of Business.* Washington: The Brookings Institution. 1942. v, 24 pp. 25 cents.

Nourse, Edwin G. *The Timing of Price Changes.* Washington: The Brookings Institution. 1943. ix, 35 pp. 25 cents.

Perkins, John A. *The Role of the Governor in Michigan in the Enactment of Appropriations.* Ann Arbor: University of Michigan Press. Michigan Governmental Studies, No. 11. 1943. ix, 197 pp. 75 cents.

Robinson, Sophia M. *Jewish Population Studies.* New York: Conference on Jewish Relations. 1943. xvi, 189 pp.

Saulnier, Raymond J., and Jacoby, Neil H. *Accounts Receivable Financing.* New York: National Bureau of Economic Research. 1943. xv, 157 pp.

Schultz, Theodore W. *Redirecting Farm Policy.* New York: The Macmillan Company. 1943. vii, 75 pp.

Shepard, E. Fenton, and Wood, William B. *The Financing of Public Schools in Michigan.* Ann Arbor: University of Michigan Press. Michigan Governmental Studies, No. 13. 1942. 139 pp. 60 cents.

Shirras, G. Findlay, and Rostas, L. *The Burden of British Taxation.* New York: The Macmillan Company. 1943. xiii, 240 pp. $3.00.

Tamagna, Frank M. *Banking and Finance in China.* New York: Institute of Pacific Relations. 1942. xxi, 400 pp. $4.00.

Van Vleck, George W. *The Panic of 1857.* New York: Columbia University Press. 1943. ix, 126 pp. $1.50.

Wells, Ralph G., and Perkins, John S. *New England Community Statistical Abstracts.* Boston: Bureau of Business Research, Boston University College of Business Administration. 1942.

Weston, John Frederick. *The Economics of Competitive Bidding in the Sale of Securities.* Chicago: The University of Chicago Press. 1943. ix, 51 pp. $1.00.

Whittlesey, Charles R. *The Banking System and War Finance.* New York: National Bureau of Economic Research. Our Economy in War, Occasional Paper 8. 1943. 53 pp. 25 cents.

Zapoleon, Marquerite W. *Community Occupational Surveys.* Washington: United States Government Printing Office. Occupational Information and Guidance Series No. 10. 1942. v, 199 pp. 25 cents.

Government Publications

Department of Commerce. *Survey of Business Research Projects at Universities.* Washington: Bureau of Foreign and Domestic Commerce. 1942. v, 188 pp.

State of New York. *Careers in New York State Government.* New York: Department of Civil Service. 1942. 62 pp.

A Economia Catarinense. Santa Catarina: Departamento Estadual de Estatistica. 1942. 36 pp.

Anuario Estadistico de la Republica Dominicana del Ano 1941. Santiago: Editorial el Diario. 795 pp.